

Communicators' Perceptions of Social Presence as a Function of Avatar Realism in Small Display Mobile Communication Devices

Sin-Hwa Kang, James H. Watt, and Sasi Kanth Ala

Social and Behavioral Research Lab
Rensselaer Polytechnic Institute
{kangs, wattj, alas}@rpi.edu

Abstract

This study describes an experiment in which 126 participants engaged via a mobile telephone simulation that included a visual display in a discussion that required self-disclosure and affective evaluation of the other participant. Participants in same gender and mixed gender dyads were represented by avatars that varied in visual realism (unmodified video, modified video, graphic display, or no visual display) and behavioral realism (static visual display versus dynamic or animated). Participants subsequently rated the Perceived Social Richness of the Medium and the Interactant Satisfaction with the conversation. Interactant Satisfaction was a new measure of social presence created to tap emotional and affective evaluations. Participants rated devices with higher-realism and more behaviorally realistic avatars as being more capable of effective social interaction, but their actual perceptions of affective dimensions of their conversational partner were essentially unaffected by visual representations.

1. Introduction

A mobile telephone is more than just a voice communication device. The small screen of a mobile telephone is increasingly being used to display various computer-generated graphics such as game characters, social agents for online shopping or education, videos, etc. The next generation of mobile telephone service promises to provide full point-to-point video communication, making the long-awaited video telephone a reality.

Previously, many studies [22,26,27,28,36] have reported on the impact of visual images of both human and nonhuman beings in virtual environments. For example, in the Human-Computer Interaction (HCI) field, a human face frequently has been used to personify computer interfaces [22] and it has been reported that the human face is one of the most attractive components in an interface [41]. Other studies [1,2,15,16,26,27,28] on the effects of avatar realism have been conducted in terms of visual realism (similarity of avatar appearance to the person communicating) or behavioral realism (physical

movement, including facial expressions, that duplicate those of the communicator) in computer mediated interactions. These studies were typically conducted with the large, and fairly high-definition, visual displays offered by full computer systems. It does not follow that these results also apply to the small screen and lower visual resolution display of mobile telephones [24].

1.1. Avatars and avatar realism

There is evidence that visual images representing communicators affect person and message perceptions in mediated environments when the images are presented on large displays [2,26,27,28]. Some results suggest that people are more engaged in virtual environments when visual images are provided, presumably because visual images augment the feeling of colocation of communicators and allow people to feel more immersed in the environment [27].

An avatar provides a means for a user to create a visual identity to support social life within virtual environments [38]. Some researchers have pointed out the importance of an avatar's anthropomorphism in the perception process, since categorization of environmental entities relies at a fundamental level on the differentiation between humans and nonhumans [28]. Avatars also provide a mechanism that communicators can use to preserve their anonymity, an important function in casual communication situations, or those in which the communicator feels vulnerable.

Previous studies [2,15,16,27,28] have shown that the use of avatars promotes a feeling of presence among users in some mediated environments. Satchell [30] has shown how users' social modes can be presented in a mobile telephone communication with a color-coded avatar. Bailenson et al. [2] describe avatars as "digital models of people that either look or behave like the people they represent" and assert that there has been no clear definition of avatars yet. They [2] consider a videoconference image (unmodified video) an avatar in some sense and claim that a videoconference is the most effective way to produce high realism in a mediated communication. In general, avatar realism is defined as possession of either visual appearance or behavioral

attributes of a human [26,27,28].

This study investigates the effect of realism of both appearance and behavior of avatars on mobile telephone communicators' perceptions of each other.

1.2. Perceptions of Interactant Satisfaction and Social Richness of Medium

The definition of social presence has been evolved from a simple binary definition of feeling or not feeling physically present to more of a continuum involving the "salience of others in mediated communication and consequent salience of their interpersonal interaction" [33].

Notably, Biocca and his colleagues [6,7] elaborated on this conception by implying psychological involvement beyond the notion of physical awareness, adding mutual understanding between interactants as an important component. They [7] suggest the use of social presence measures to evaluate the performance of mobile and wireless telecommunication.

Biocca et al. [7] describe "subjective social richness of the medium" as one measurement of social presence, an idea which originates in Short et al.'s definition [33]. This measures users' attitudes towards the medium, rather than their awareness of the other communicator. Daft and Lengel [10,11] propose that a rich medium with a high bandwidth provides many cues, while a lean medium has low cue delivery because of low bandwidth. The subjective social presence is likely to covary with this idea of richness of medium.

Biocca et al. [7] also introduced social attraction as an element of social presence, defining it as "social judgments of the other." Nowak [26,28] used social attraction and credibility as components of interactant satisfaction in virtual environments while exploring the effect of anthropomorphism and agency in task-oriented virtual environments. In these task-oriented studies, Nowak measured interactants' knowledge of a particular subject to establish credibility of interaction partners. She [27] also used measures of the social richness of medium to indicate users' sense of social presence in virtual environments.

Nowak [26] has described how people are likely to use visible indicators such as a face to reduce uncertainty between strangers, and also to evaluate credibility of other interactants within mediated or unmediated environments. It is hypothesized that if people have no visible images of other interactants in mediated environments, they construct a mental image of the other interactant that emphasizes the similarities between themselves and the other in order to decrease feelings of uncertainty [26,27,42].

Kang, Watt, and Isbister [19] used measurements of perceived social richness of medium and interactant satisfaction to evaluate users' perceptions of each other when the communicators were represented by static anthropomorphic avatars. Other studies [2,16] investigated the impact of both appearance and behavior of avatars on users' copresence (also referred to as social presence) in computer mediated interactions.

1.3. Social Use of Mobile Telephone Communication

Previous studies [3,8,25,32] have shown that people use computer-mediated communication (CMC) to strengthen social ties, for instance by exchanging emails or instant messages. People are likely to use a mobile telephone to make social ties stronger [17,21,23]. Mobile telephone cameras have allowed people to share pictures and other visual images which may help individuals construct personal and group memory [40]. Studies of mobile computing in Europe have shown strangers share their personal information or stories via mobile telephones [20,29]. Mobile social software (MoSoSo) is increasingly used for communication between strangers [13,34]. Tomita [39] describes how a mobile telephone may be used for social tie-shaping purpose between strangers. Humphrey [17] explores how a mobile telephone serves to foster new social bonds among urban dwellers, including both people who are known to the user and strangers. She reports that MoSoSo applications help to create different kinds of social networking for urban living. Overall, researchers demonstrate that MoSoSo applications facilitate development of social ties among strangers [13,17,34].

1.4. Measuring Users' Social Presence in Emotionally Engaged Mobile Phone Communication

In mediated social interaction that involves visual information, it is important that a medium is able to convey people's emotional expressions through their nonverbal behaviors to make emotionally engaged communication more effective in establishing social ties. This kind of communication function differs from the task-oriented functions typically studied in CMC research. A different type of measurement is required to evaluate this more affective type of interactant satisfaction.

In task-oriented communication, which has been the norm in CMC studies that typically come with an implied organizational or design project objective, researchers have usually focused on measuring interactants' knowledge or performance, rather than their emotional or affective perceptions. Likewise, most CMC studies have

been conducted on computer systems having large visual displays that produce a visual experience that is very different from that possible with a very small mobile telephone screen. Both these limitations imply the need for research into the impact of avatars in communications whose utility is strongly or primarily emotional, via technology which permits only very limited visual displays.

Kang, Watt, and Isbister [19] describe development of a measure of perceived social richness of medium and an interactant satisfaction measure based on users' emotional perceptions of each other. In this study, these measures are used to evaluate the impact of the realism of avatars when they are used in mobile telephone communication that is not primarily task-oriented.

2. Research Problem and Questions

Previous studies [1,2,15,16,26,27,28] have explored some effects of avatar visual or behavioral realism in computer mediated interactions. However, few studies of avatars have investigated the impact of avatars' behavioral realism on users' social or emotional responses. Some studies [1,2,15,27] conclude that users' social presence is lower when they experience a large incongruity between the behavioral realism and the visual realism of symbolic social actors (agents or avatars).

Garau and colleagues [16] demonstrate that avatars' higher visual realism requires higher behavioral realism to increase users' sense of copresence (also referred to as social presence).

In this research the behavioral realism of avatars' facial expressions is examined. Facial expression in communication is a primary way that humans transmit emotional signals [14,15]. Thus the behavioral realism of avatars' facial movements may have an impact on the social presence that they induce.

The primary research questions addressed in this research concern the effects of the visual and behavioral realism of avatars on users' levels of social presence when they are used in a medium which is typically used for audio-only communication, and which has limited graphic abilities.

Studies of social presence in computer-mediated settings often conclude that the highest presence is obtained in face-to-face communications which are accompanied by rich nonverbal cues [33]. Media richness theory [10,11,12] claims that richer media can facilitate emotionally arousing tasks that include high equivocality. However, some studies report that task-oriented groups report less sense of social presence when time limits for interactions are imposed [42,44]. In fact, Walther's "hyperpersonal communication" [42] theory supports the idea that people feel more affection and emotion with their interactants when they have fewer cues combined

with more frequent and longer duration interactions. This leads to the prediction that "lean" media may actually provide a stronger perception of emotional social presence in this case than do "rich" media. With this view, Walther et al. [43] challenge the idea that increased visual fidelity automatically increases the sense of social presence. By extension, this view may also extend to behavioral realism of avatars, too.

In their study of avatars as a visual addition to mobile telephone communication [19], Kang, Watt and Isbister report that static images of avatars with different degrees of visual anthropomorphism (visual similarity to humans) do not have any impact on users' perceptions of interactants and medium during communication that is primarily emotional and non-task-oriented. They found that participants felt more engaged in their communication when they did not use any anthropomorphic images representing their partners. The use of static anthropomorphic images had no significant impact on the perception of the social richness of the mobile telephone medium. Some participants in their study who looked at the high anthropomorphic image reported that they expected a more animated or real image, and some participants who looked at the low anthropomorphic avatar reported that they expected a more anthropomorphic image. These expectations of the avatar image quality and behavior may have affected their evaluations of the interaction partner and the medium and produced the lower engagement reports. It is possible that dynamic, more realistic images might provide greater emotional connection between conversational partners and could reverse this evaluation.

However, mobile telephone users may prefer that these images do not make them identifiable to everyone with whom they communicate. Visual and behavioral realism are not necessarily synonymous with faithful video images of the communicator. Video can be modified to mask identity [18], or a graphic avatar may be substituted for a picture of those who want to maintain their privacy and be released from "social norms of politeness" [15] to express their feelings and emotions more freely.

But, the impact of these visual modifications or replacements is not clear. They will likely convey fewer cues (nonverbal information critical to effective emotional communication) compared to unmodified video. If the "rich media" hypothesis is correct, this should reduce interactants' satisfaction with the communication and with the medium. But the "hyperpersonal communication" hypothesis predicts the opposite, at least in the situation of long interactions. In this case, the outcome is best stated as open research questions rather than as hypotheses:

RQ1: Do people report **higher interactant satisfaction** in more emotionally engaged communication when their interaction partners are

represented by modified video or animated graphics avatars, as compared to unmodified video?

RQ2: Do people report **higher perceived social richness of medium** in more emotionally engaged communication when their interaction partners are represented by modified video or animated graphics avatars, as compared to unmodified video?

Bickmore and Picard [5] report higher user ratings for video, as opposed to audio-only, mediated communication only for social tasks. The availability of visual nonverbal communication plays an important role in this result.

The implication of the cited studies is that unmodified video, behaviorally correct modified video or graphic animated avatars should all provide a richer emotional exchange than static avatars or no avatars when the time of interaction is limited. Furthermore, static avatars should provide richer cues than audio-only communication, when the time of interaction is limited. This leads to these hypotheses

H1: People will report **higher interactant satisfaction** when their interaction partners are represented by dynamic (high behavioral realism) avatars, compared to a communication in which there are no avatars.

H2: People will report **higher interactant satisfaction** when their interaction partners are represented by static (low behavioral realism) avatars, compared to a communication in which there are no avatars.

H3: People will report **higher perceived social richness of medium** when their interaction partners are represented by dynamic (high behavioral realism) avatars, compared to a communication in which there are no avatars.

H4: People will report **higher perceived social richness of medium** when their interaction partners are represented by static (low behavioral realism) avatars, compared to a communication in which there are no avatars.

3. Experimental Design

The basic experimental design is a 3 (visual realism) x 2 (behavioral realism) x 3 (gender dyads) factorial between-subjects experiment involving three factors: i) three levels of visual realism of avatars: an unmodified video, a modified video, and a graphic avatar; ii) two levels of behavioral realism of avatars - a high realism (dynamic visual) and a low realism (static visual); iii) three levels of type of gender dyad – male-male, male-female, and female-female.

As the focus of this study is on non-task-oriented communication in which the goal is to get to know the other person, the type of gender dyad was included as a control. Especially in the case of mixed gender dyads, the interactant satisfaction may be affected by gender differences in communication style, willingness to self-disclose, etc. While these differences are not central to the research questions being investigated here, explicitly including gender dyad effects in the analysis controls for these possibilities.

In the low behavioral realism conditions, subjects see a static image during a simulated mobile telephone conversation. In the high behavioral realism conditions, the subjects see video or dynamic animations. To test the hypotheses contrasting the presence and absence of avatars of any type, an additional audio-only condition with 9 dyads was added to the design. The rationales for the choices of the visual realism conditions are outlined below.

Each experimental condition in this research was presented to three gender combinations: male-male, male-female, and female-female with three replications of each type of dyad. Avatars were displayed on the mock-up of a standard mobile telephone screen on a laptop computer, with screen size of the mockup similar to actual mobile telephone screens. The experiment measured a number of dimensions of social presence, not all of which are reported here. Measures used validated Likert-type scales [19,26,27], which are described below.

3.1. Participants

One hundred twenty six participants (63 males and 63 females) from a northeastern university were recruited to participate in the experiment (see Table 1).

Table 1. Numbers of paired participants in seven experimental conditions.

	Unmodified Video	Modified Video	Graphic Avatar	No Images (Audio-only)
Dynamic Status	18 (3 gender dyads x 3 replications)	18 (3 gender dyads x 3 replications)	18 (3 gender dyads x 3 replications)	18 (3 gender dyads x 3 replications)
Static Status	18 (3 gender dyads x 3 replications)	18 (3 gender dyads x 3 replications)	18 (3 gender dyads x 3 replications)	

3.2. Procedure

The participants were randomly assigned to one of the visual realism and behavioral realism experimental conditions shown in Table 1. Each participant was paired with someone whom they were unlikely to know beforehand to create the gender dyads. The interaction took place in two separate rooms where the paired participants were placed at different times, to avoid any initial face-to-face interaction.

Participants were given a hypothetical conversational scenario where each assumed the role of a student who sought to find out if the other person was a suitable match with whom an apartment could be shared and with whom s/he could be friendly. This conversational topic was chosen to stimulate investigation of the other person's personality and emotional status. Interactants needed to encourage the other person to self-disclose, and had to provide potentially sensitive information about themselves.

Anonymity in this kind of self-revealing interaction is likely to be important in actual conversation, so the use of avatars to mask visual identity was a natural situation. The typical conversation lasted about ten minutes. Participants were asked to complete a post-test questionnaire with measurements that operationalized the variables described below.

3.3. Apparatus

Avatar designs for the different experimental conditions are shown in Figure 1. Avatars included an unmodified video image of the participant representing the highest visual realism with very high behavioral realism in the dynamic avatar conditions (Figure 1a); a modified video that was lower on visual realism to produce identity masking, but also high on behavioral realism in the dynamic avatar conditions (Figure 1b); and graphic avatars that were matched to the participants' gender, representing low visual realism, with high behavioral realism in the dynamic avatar conditions produced via animation that tracked the participant's facial expressions (Figure 1c and 1d). In the low behavioral realism conditions, the avatars were simply still visuals. The unmodified and modified video avatars in these conditions were created by capturing a single frame of video of the participant and displaying it for the duration of the experiment.

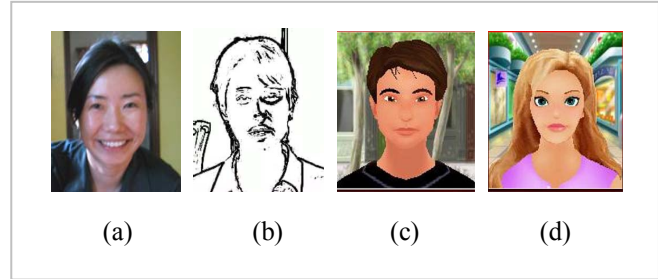


Figure 1. (a) an unmodified video; (b) a modified video; (c) a graphic male avatar; (d) a graphic female avatar.

To create the modified video condition, the findings of Zhao and Stasko [45] were applied. They evaluated four types of video-filtering techniques designed to provide participants' identity masking and reported that an "edge-detector" filtered image like Figure 1b provided lower actor identification than a pixelized one commonly used for masking identity in news broadcasts [9]. In a pilot test conducted as a preliminary to this research, seventeen persons were asked to identify from a set of seventeen photographs the image of a person presented in a short video that had been modified by this technique. Only three persons could do so, even though the edge detection preserved critical behavioral and nonverbal information in the video.

Finally, for the animated avatars, the commercial product "Logitech Video Effects" was used. This software incorporates proprietary NevenVision technology to create animated avatars that exhibit realistic facial expressions based on transferring motion from real-time tracking of 22 points on a human face within the frame of a Webcam.

To simulate communication with a mobile telephone equipped for real-time video, a mobile telephone mock-up with a realistic small-size screen (2.7") and display resolution (240 x 320 pixels) was created and displayed on a laptop screen (see Figure 2). The mock-up was written in Java using Java Media Framework (JMF). This allowed real-time video processing for the modified video condition. To create the animated avatar, a Webcam (Logitech QuickCam Orbit MP equipped with face-tracking function) captured users' facial expressions. A commercial hands-free mobile telephone headset connected to the laptop was provided to users for the audio conversation. The audio and video were combined and transmitted over a high bandwidth Local Area Network connecting the two rooms in which the participants were located.

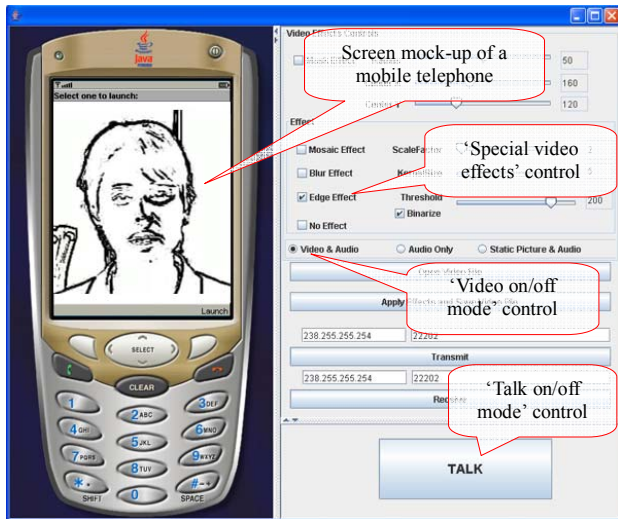


Figure 2. Mock-up of a mobile telephone & control panel of video/audio transmission. During the experimental session, the control panel was not visible.

3.4. Measurement Instruments

Two types of social presence were operationalized to measure users' perceptions of the other interactant and of the medium.

3.4.1. Interactant Satisfaction. A fifteen item Likert-type scale with an eight point metric (1 = Strongly Disagree; 8 = Strongly Agree) was created. Six of the items were adopted from the items of the social attraction scale used in Nowak's study [26]. The other nine items, called emotional credibility, were newly created to measure interactants' emotional perceptions of their interaction partners [19] (see Table 2). The new items were derived from the evaluation dimensions of emotional intelligence test [35] which originated in the studies of Mayer and Salovey [31].

Table 2. Items used in the measurements of Interactant Satisfaction.

Interactant Satisfaction
<i>Social Attraction Items</i>
I think s/he could be a friend of mine.
I would like to have a friendly chat with her/him.
We could never establish a personal friendship with each other.
S/he just wouldn't fit into my circle of friends.
S/he would be pleasant to be with.
I don't care if I ever get to interact with her/him again.
<i>Emotional Credibility Items</i>
S/he recognizes my feelings and emotions.
S/he expresses feelings and emotions appropriately for the situation.
S/he uses feelings and emotions to create or organize thinking.
S/he uses feelings and emotions to make a decision or judgment.

S/he uses feelings and emotions to facilitate problem solving and creativity.
S/he responds appropriately to positive and negative emotions.
S/he understands complex feelings.
S/he knows how to control her/his own feelings and emotions effectively.
S/he handles others' feelings and emotions sensitively and effectively.

3.4.2. Perceived Social Richness of Medium. A five item Likert-type scale was created with an eight point metric for items (1 = Very Little; 8 = Very Much or 1 = Very Unlikely; 8 = Very Likely). The five items were modified from the items of "subjective social richness of the medium" scale used in Nowak and Biocca's study [27]. The items were worded to be applicable to a mobile telephone communication (see Table 3).

Table 3. Items used in the measurements of perceived social richness of medium.

Perceived Social Richness of Medium
To what extent did you feel able to assess your partner's reactions to what you said?
To what extent was this like a face-to-face meeting?
To what extent was this like you were in the same room with your partner?
How likely is it that you would choose to use a mobile telephone interaction for a meeting in which you wanted to persuade other person of something?
To what extent did you feel you could get to know someone that you met only through a mobile telephone?

3.4.3. Gender Dyad Control Variable. In the data analysis, gender combinations of interactants were considered a measured control variable. Participant dyads were evenly distributed across male-male (21), male-female (21), and female-female (21) pairs. Any gender main effects due to interaction style or similar effects are thus controlled, as are interaction effects between the type of avatar and gender pairings.

4. Results

Reliability tests showed good internal consistency among the fifteen items of Interactant Satisfaction measure when they were combined into a summed scale (Cronbach's alpha = .90), as well as the five items of the Perceived Social Richness of Medium measure (Cronbach's alpha = .83).

A between-subjects ANOVA was conducted to explore the effects of avatar visual and behavioral realism as well as mixed-gender dyad type on levels of Interactant Satisfaction and Perceived Social Richness of Medium.

4.1. Results for Interactant Satisfaction

The first research question (RQ1) concerned the impact of avatar realism on the Interactant Satisfaction measure.

Overall ANOVA results did not show a statistically significant main effect for realism of visual appearance (unmodified video, modified video, or graphical avatar), behavioral realism (dynamic/animated or static), or the type of gender dyads (M-M, M-F, or F-F) on this variable (all F-values had $p > .05$). Interaction effects were also not significant. There was no indication that visual representations of the communicators resulted in systematic differences in the primarily emotional and affective perceptions of participants of their interaction partner.

Hypothesis 1 predicted that participants would report higher Interactant Satisfaction when dynamic avatars were used, as compared to no avatars (audio-only). This hypothesis was not supported.

There were no statistically significant main effects on Interactant Satisfaction for the type of avatar or the type of gender dyad. All interaction effects also did not reach statistical significance ($p > .05$ for all main and interaction effects). The presence of avatars did not produce higher levels of Interactant Satisfaction than did simple audio communication.

Hypothesis 2 predicted that participants would report higher Interactant Satisfaction when static avatars were used, as compared to audio-only communication. This hypothesis was not supported. There was no significant main effect for the visual appearance of the static avatar.

However, the results did reveal a statistically significant main effect for the type of gender dyad [$F(2, 60)=3.32, p=.043$] on Interactant Satisfaction. Post-hoc comparisons using the Tukey HSD test showed that the mean score for the 'male-male' dyads ($M=5.78$) and 'female-female' dyads ($M=5.70$) significantly differed from the 'male-female' dyads ($M=5.18$).

This indicates that participants were more satisfied with the conversation when interacting with the same gender partners when the effects of the presence or absence of static avatars were controlled. However, given that the low level of significance and the lack of a similar effect when dynamic avatars were examined, this finding must be interpreted with caution.

Interaction effects in this analysis did not reach statistical significance for the type of visual appearance, the type of behavior and the type of gender dyad ($p > .05$ for all interactions).

4.2. Results for Perceived Social Richness of Medium

The second Research Question involved the impact of

avatar realism on the Perceived Social Richness of Medium variable.

Three-way ANOVA (visual realism by behavioral realism by gender dyad) results did reveal a statistically significant main effect for the visual appearance of the avatar [$F(2, 90)=6.11, p=.003, \eta^2=.12$] on Perceived Social Richness of Medium. The eta-square value could be considered a large effect size in Cohen's (1988) terms. Post-hoc comparisons using the Tukey HSD test showed that the mean score for the graphic avatars condition ($M=4.08, SD=1.04$) was significantly lower on Perceived Social Richness of Medium than was the mean for unmodified video avatars ($M=4.84, SD=1.40$), and modified video avatars ($M=5.03, SD=1.33$), which were not significantly different from each other.

The results also showed a statistically significant and large main effect for avatar behavioral realism [$F(1, 90)=15.42, p=.000, \eta^2=.15$] on Perceived Social Richness of Medium (see Table 4). The medium was rated higher when avatars exhibiting high behavioral realism (dynamic avatars) were used, regardless of the visual appearance of the avatar.

Table 4. Effects of Avatar Differences on Perceived Social Richness of Medium.

<i>Type of Avatar</i>	<i>Avatar Dynamic Behavior</i>	<i>Avatar Static Behavior</i>
Unmodified Video	M=5.38, SD=1.43	M=4.31, SD=1.18
Modified Video	M=5.52, SD=1.25	M=4.53, SD=1.26
Graphic	M=4.43, SD=1.00	M=3.73, SD=.98

The results did not reveal a statistically significant main effect for gender dyads or for interaction effects ($p > .05$ for all effects).

Hypothesis 3 predicted that using dynamic avatars would result in higher Perceived Social Richness of Medium when compared to no avatars (audio-only).

Two-way ANOVA results revealed a statistically significant and large main effect for the visual appearance (unmodified video, modified video, graphic, no avatar) [$F(3, 60)=6.03, p=.001, \eta^2=.23$] on Perceived Social Richness of Medium. Hypothesis 3 is supported.

Post-hoc comparisons using the Tukey HSD test indicated that the mean score for the no avatars condition ($M=3.83, SD=1.67$) and the graphic avatars condition ($M=4.44, SD=1.00$) were significantly lower than the unmodified video avatars condition ($M=5.38, SD=1.43$), and the modified video condition ($M=5.52, SD=1.25$).

The results did not show a statistically significant main effect for the type of gender dyad or for interaction effects ($p > .05$ in all cases).

Hypothesis 4 predicted that participants would report a higher Perceived Social Richness of Medium when static avatars were used, as compared to no avatars. Hypothesis 4 was not supported.

ANOVA results did not reveal a statistically significant main effect either for the type of visual appearance or for the type of mixed-gender dyad, and no interaction effects were significant ($p > .05$ for all effects).

5. Conclusions and Discussion

This study differs from the previous work in examining primarily non-task oriented communication that requires emotionally-sensitive interaction, conducted in a medium that provides a much smaller visual display with limited resolution. Previous studies [4,26,27, 28,37,38] have reported that people feel more immersed in virtual environments when they interact using visual images, possibly because they are able to reduce their uncertainty about a partner whom they have not met before and do not know [26,28].

The results of this study provide some qualifications to this conclusion. Participants perceived the medium they were using to be more *capable* of social interaction when avatars were used, but their perceptions along affective dimensions of the other party with whom they had just communicated were remarkably unaffected by visual representations.

This may have been the result of kind of emotional and affective communication that was occurring, which may be less reliant on visual information, as opposed to the more pragmatic tasks addressed in prior studies, or it may have been a function of the small visual display's limited abilities to convey nonverbal cues. In any event, there is a discrepancy between the perception of the medium by users (it was seen as more capable when visuals were present than when they were absent, and even more so with highly realistic visuals) and the actual outcome of the communication (emotional satisfaction was not improved by the addition of avatars, regardless of their realism).

Past studies [42,43,44] have concluded that people who looked at static images of their conversational partner rated the interaction higher on social presence (affection and social attraction) after short-term interactions, while the opposite was true for longer-term interaction. Neither finding was replicated here, and again, it may have been a function of the type of communication task and the technology used.

Although the study has some limitations (outlined below), if its results can be assumed to hold generally, they raise questions about the value of providing visual avatars on small mobile telephone screens when the primary goal is facilitating social communication.

There is clearly marketing value to providing a visual representation of the communicator. Respondents perceived the simulated telephones with modified video (for those who would like to preserve anonymity) as the most suited for social communication, followed by

unmodified video, then graphic avatars, with the current audio-only function being seen as least capable. It is likely that they would choose to purchase and use mobile devices with graphics capabilities rather than audio-only devices as a result of this perception.

However, this is different from saying that such devices facilitate effective social communication. There was no evidence found in this study of more effective (from the standpoint of the communicators) emotional outcomes when graphics were present.

It is interesting to note that participants who used modified video, as would be the case if they were trying to protect their visual identity, found the medium to be nearly as socially rich as unmodified video, with graphic avatars trailing significantly in this respect. Again, if the results of this study can be generalized, providing mobile telephone users with a video processing option that preserves anonymity seems to be preferable to providing them with a graphical avatar, even if the avatar is animated and preserves some of the facial expressions of the user.

5.1. Limitations

This study was aimed particularly at examining communication between strangers in a single emotionally engaged situation. This is only one type of communication, and it is possible that these results are limited to that situation.

The fixed communication task used in this experiment also limits the external validity of the findings somewhat. The requirement that participants assume a role of a searcher for a roommate may not have produced the kind of engagement that a real situation would impose.

Presenting a telephone video screen mockup on a laptop screen, even though it was the same size and resolution as an actual device, might have generated different reactions in participants. To partially compensate for this, the participants were placed within arms-length of the laptop, at a distance that would normally be used in mobile telephone use. However, an actual telephone could be moved closer to the user's eyes, while the laptop setup did not permit this.

The animation technology used in this experiment was not able to reproduce eye gaze movements for graphic avatars. This might have contributed to users' lower ratings for those avatars, as compared with the videos.

Finally, the subjects of this experiment, like those of many other CMC studies, were college students. While the conversational task was chosen to be relevant to this group, there is no denying that college students are hardly representative of the general population.

5.2. Future Work

The findings of this study pose an important unanswered research question: are they due to the nature of the communication task (affective rather than pragmatic), or to the limitations of the small visual display? Further research is needed to disentangle the two effects.

The range of visual characteristics of avatars needs further extension, as well. This includes studies that examine the usefulness of full-body avatar kinesics rather than being restricted to dynamic facial appearances.

There are also two larger research issues embedded in studies of mediated communication like this one that involve both pragmatic and affective communication.

The first is some systematic way to describe the impact of the communication task itself. Prior research has focused on individuals and technologies, with the communication task usually being an unexamined constant. But conflicting results and failures to replicate findings in prior research are possibly due to differences in the communication settings and motives imposed by the study protocols.

It is important to be able to distinguish between the effects of different mediating technologies found in a purely pragmatic communication and those found in a purely affective communication, or between communication tasks that are more concrete or visual and those that are more abstract, to mention two examples. Relationships and effects are likely to vary in these differing situations. We currently have no useful taxonomy to identify and define these kinds of tasks, nor research to describe their expected impact on communication outcomes.

The communication outcomes themselves need to be examined in more elaborated experimental designs that include other measurements of social presence. This study has made a small contribution to this task by introducing emotional credibility as a measured construct that extends the definition of social presence. Other such extensions are needed.

Acknowledgements

We would like to thank Dr. Katherine Isbister, Dr. Robert Krull, and Dr. Linnda Caporael for their advice on this research.

References

[1] Bailenson, J. N., Blascovich, J., Beall, A. C., & Loomis, J. Equilibrium revisited: Mutual gaze and personal space in virtual environments. *Presence: Teleoperators and Virtual Environments*, 10, 583–598 (2001).

[2] Bailenson, J.N., Yee, N., Merget, D., & Schroeder, R. The Effect of Behavioral Realism and Form Realism of Real-Time Avatar Faces on Verbal Disclosure, Nonverbal Disclosure, Emotion Recognition, and Copresence in Dyadic Interaction. *PRESENCE: Teleoperators and Virtual Environments* (2006).

[3] Barnes, S. *Computer-Mediated Communication: Human-to-human communication across the Internet*. Allyn and Bacon. (2003).

[4] Benford, S., Greenhalgh, C., Rodden, T., & Pycock, J. To what extent is cyberspace really a space? Collaborative virtual environments. *Communications of the ACM*, 4(7). 2001.

[5] Bickmore, T. & Picard, R. Subtle Expressivity by Relational Agents. *Proc. CHI'03 Workshop on Subtle Expressivity for Characters and Robots* (2003)
<http://affect.media.mit.edu/pdfs/03.bickmore-picard.pdf>.

[6] Biocca, F. & Harms, C. What is social presence? In F. Gouveia & F. Biocca (Eds.), *Proc. Presence 2002*. University of Fernando Pessoa Press (2002).

[7] Biocca, F., Harms, C., & Burgoon, J. K. Toward a more robust theory and measure of social presence: Review and suggested criteria. *Presence: Teleoperators and Virtual Environments*, Vol. 12, No. 5 (2003).

[8] Boneva, B. & Kraut, R. Email, gender, and personal relationships. In Wellman, B. and Haythornthwaite, C. (eds.) *The Internet in everyday life*. Blackwell Publishing Ltd. (2002).

[9] Boyle, M., Edwards, C., & Greenberg, S. The effects of filtered video on awareness and privacy. *Proc. Conference on Computer Supported Cooperative Work (CSCW '00)*. [CHI Letters 2(3)], ACM Press, 1–10 (2000).

[10] Daft, R. L. & Lengel, R. H. Information Richness: A new approach to managerial behavior and organization design. *Research in Organizational Behavior*, 6 (1984).

[11] Daft, R. L. & Lengel, R. H. Organizational information requirement, media richness and structural design. *Management Science*, 32 (1986).

[12] Dennis, A. & Kinney, S. Testing Media Richness Theory Information Systems Research: The Effects of Cues, Feedback, and Task Equivocality. *Information Systems Research*. Vol. 9, No. 3 (1998).

[13] Eagle, N. & Pentland, A. Social serendipity: Mobilizing social software. *IEEE Pervasive Computing*, 28-34, April-June 2005.

[14] Ekman, P. & Friesen, W. The repertoire of nonverbal behavior: Categories, origins, and coding. *Semiotica* (1969),1, 49-98.

[15] Garau, M. *The Impact of Avatar Fidelity on Social Interaction in Virtual Environments*. Ph.D. thesis, Department of Computer Science, University College London (2003).

[16] Garau, M., Slater, M., Vinayagamoorthy, V., Brogny, A., Steed, A., & Sasse, M. A. The impact of avatar realism and eye gaze control on perceived quality of communication in a shared immersive virtual environment. *Proc. CHI'03*. ACM Press (2003).

[17] Humphreys, L. Mobile Devices and Social Networking. *Pre-Conference, International Communication Association*, (2006).

- [18] Kang, S. The impact of digital iconic realism on anonymous interactants' mobile phone communication. *Proc. CHI'07*. ACM Press (2007).
- [19] Kang, S., Watt, J., & Isbister, K. The effect of static anthropomorphic images on emotion perception in mobile-phone communication. *Proc. 9th International Workshop on Presence* (2006).
- [20] Jungnickel, K. Urban Tapestries: Sensing the city and other stories. *Proboscis Cultural Snapshots Number Eight* (2004). <http://proboscis.org.uk>.
- [21] Katz, J. & Aakhus, M. (eds.) *Perpetual Contact*. Cambridge University Press (2002)
- [22] Koda, T. Agents with faces: The effects of personification. *Proc. HCI 1996* (1996).
- [23] Ling, R. *The mobile connection: The cell phone's impact on society*. Morgan Kaufmann Publishers (2004).
- [24] Mosmondor, M., Kostucic, T., & Pandzic, I. LiveMail: Personalized avatars for mobile entertainment. *Proc. 3rd International Conference on Mobile Systems, Applications, and Services* (2005).
- [25] Norman, N., Sunshine, H., & Erbring, L. Internet use, interpersonal relations, and sociability: A time diary study. In Wellman, B. and Haythornthwaite, C. (eds.) *The Internet in everyday life*. Blackwell Publishing Ltd. (2002).
- [26] Nowak, K. L. The Influence of Anthropomorphism and Agency on Social Judgment in Virtual Environments. *Journal of Computer Mediated Communication*, 9(2), (2004). <http://jcmc.indiana.edu/vol9/issue2/nowak.html>.
- [27] Nowak, K. & Biocca, F. The effect of the agency and anthropomorphism on users' sense of telepresence, copresence, and social presence in virtual environments. *Presence: Teleoperators and Virtual Environments*, Vol. 12, No. 5 (2003).
- [28] Nowak, K. & Rauh, C. The influence of the avatar on online perceptions of anthropomorphism, androgyny, credibility, homophily, and attraction. *Journal of Computer-Mediated Communication*, 11(1), (2005). <http://jcmc.indiana.edu/vol11/issue1/nowak.html>.
- [29] Persson, P. Nokia Sensor (2005). http://press.nokia.com/PR/200506/997857_5.html
- [30] Satchell, C. "Contextualising Mobile Presence with Digital Images" In PICS Workshop, *Ubicomp* (2006).
- [31] Salovey, P., Brackett, M., & Mayer, J. *Emotional Intelligence: Key readings on the Mayer and Salovey Model*. Dude Publishing (2004).
- [32] Shiu, E. & Lenhart, A. How Americans use instant messaging. *PEW INTERNET & AMERICAN LIFE PROJECT* (2004). <http://www.pewinternet.org/>.
- [33] Short, J., Williams, E., & Christie, B. *The social psychology of telecommunication*. London: John Wiley & Sons (1976).
- [34] Smith, I. E. Social mobile applications. *IEEE Computer Society*, 38(4):84-85, April 2005
- [35] Smith, J. Emotional Intelligence Report: EI16 Questionnaire (2004). http://www.myskillsprofile.com/eiq16_sample_report.pdf
- [36] Takeuchi, A. & Naito, T. Situated facial displays: Towards social interaction. *Proc. CHI '95 Human Factors in Computing Systems*, Addison Wesley (1995).
- [37] Talamo, A. & Ligorio, B. Strategic identities in cyberspace. *CyberPsychology & Behavior*, 4(1). 2001.
- [38] Taylor, T. L. Living digitally: Embodiment in virtual worlds. In R. Schroeder (Ed.), *The Social Life of Avatars: Presence and Interaction in Shared Virtual Environments*. London: Springer-Verlag (2002).
- [39] Tomita, H. Geitai and the Intimate Stranger. In Ito, M., Okabe, D. & Matsuda M. (eds.) *Personal, Portable, Pedestrian*. The MIT Press (2005).
- [40] Van House, N., Davis, M., Takhteyev, Y., Ames, M., & Finn, M. The social uses of personal photography: Methods for projecting future imaging applications. (2004). http://www.sims.berkeley.edu/~vanhouse/photo_project/pubs/vanhouse_et_al_2004b.pdf.
- [41] Walker, J. H., Sproull, L., & Subramani, R. Using a human face in an interface. *Proc. CHI '94 Human Factors in Computing Systems*, Addison Wesley (1994).
- [42] Walther, J. B. Computer-mediated communication: Impersonal, interpersonal, and hyperpersonal interaction. *Communication Research*, 23, 3-43 (1996).
- [43] Walther, J., Slovacek, C. & Tidwell, L. Is a picture worth a thousand words? *Communication Research*, Vol. 28 (1). (2001).
- [44] Walther, J. B. & Parks, M. R. Cues filtered out, cues filtered in: Computer-mediated communication and relationships. In I. M. L. Knapp & J. A. Daly (Eds.), *Handbook of interpersonal communication* (3rd ed., pp. 529- 563). Thousand Oaks, CA: Sage (2002).
- [45] Zhao, Q. & Stasko, J. Evaluating image filtering based techniques in media space applications. *Proc. 1998 ACM conference on Computer supported cooperative work* (1998).