Social Copresence in Anonymous Social Interactions Using a Mobile Video Telephone

Sin-Hwa Kang, James H. Watt, and Sasi Kanth Ala
Social and Behavioral Research Laboratory
Rensselaer Polytechnic Institute
Troy, NY 12180-3590 USA
{kangs, wattj, alas}@rpi.edu

ABSTRACT
In this paper, we describe research exploring the effect of behavioral and visual realism of avatars on users' social copresence in emotionally engaged conversations conducted via a simulated mobile video telephone. We offer an elaborated definition of Social Copresence to better measure users' engagement with conversational partners in social interactions that do not involve specific tasks or concrete outcomes. We investigate ways to secure mobile telephone users' anonymity while preserving their most important nonverbal affective behaviors. Experimental results with 180 participants using different combinations of static and dynamic, high and low iconic (both video and graphically animated) avatars show increased Social Copresence with dynamic high-iconic (similar to the human communicator) avatars incorporating correct facial expressions, even when these are presented on the small screen of mobile telephones in such a way that individual identities are masked. The results point to an economical combination of behavioral and iconic realism of avatars that produces maximum emotional engagement in anonymous social interactions using mobile video telephones.

Author Keywords
Mobile phone communication, avatar realism, anonymity, affective behavior, social presence, social copresence.

ACM Classification Keywords
H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous; J.4 Computer Applications: Social and Behavioral Sciences---Psychology.

INTRODUCTION

With the introduction of broadband wireless telecommunications services that support mobile video viewing, and products like the Apple iPhone and the Motorola Ojo that combine personal communications with video capabilities, the prospect of the long-awaited full point-to-point personal video telephone system seems imminent. But the introduction of a moving visual display coupled with point-to-point voice service carries a wide array of questions about the impact of such capability on social interactions.

CHI 2008, April 5–10, 2008, Florence, Italy.

Social Connection and Mobile Communication

Smith and Mackie [38] state that humans' social bonds are elicited by their basic need to be connected with others, which can be called social connectedness. Social connectedness is defined as the presence of relationships among people which are essential to their health and successful life and are fostered when those relationships are maintained constructively [23].

Recent studies [7,8,13,36,40] address the importance of mobile telephone use in fostering social connectedness. Mobile communication technologies provide an easy means of strengthening social bonds among friends, while also making contact and social bonding easier among strangers. Mobile communication has allowed people to make dynamic plans and construct fluid social networks through functional and expressive management [20]. This in turn has moved people's communication patterns into a "social-network-based paradigm," where people shape their bonds based on mutual trust, rather than being governed by their physical location [29].

In the research reported here, we explore an important aspect of social bonding, namely the ability of the communication medium to produce a sense of "social copresence," which will be defined in this study, between two communicators. This sense of being co-connected is likely to be affected by the addition of a visual element to voice communication, as visuals can provide important nonverbal visual signals that may produce the sense of being jointly present and having access to each other's thoughts.

But there are situations where a communicator may wish to foster copresence in order to have a more satisfying and effective communication, but at the same time may not wish to share his or her identity with a conversational partner. This might be the case when first meeting

a stranger or when conducting casual business via mobile technologies (e.g., searching for a plumber). Including nonverbal visual information in these cases to improve communication effectiveness is desirable, but permitting visual recognition of the communicator is not desirable. In this research, we will investigate methods of preserving visual anonymity while retaining nonverbal information.

Importance of Nonverbal Communication in Mediated Interactions

Ekman and Friesen [9] demonstrate that nonverbal behavior is used as a primary channel to communicate emotion and to indicate changes in "the quality of an ongoing interpersonal relationship." Garau [11] describes "conversation management and the communication of emotion" as two critical functions of nonverbal behaviors in face-to-face communication. She addresses the importance of emotion, which is vital in the communication of understanding and which can be captured through monitoring people's facial expression, eye gaze, and body movements.

Littlejohn [21] contends that appropriate nonverbal behavior produces effective communication, as facial expressions can provide cues that eliminate conversational confusion. It has been shown that the facial cues convey emotional signals most efficiently [10,11,14,30]. Izard [14] addresses the value of facial expression specifically to describe its critical role in conveying emotion signals in a social relationship. Poggi and Pelachaud [30] also demonstrate that people show their social attitude or relationship through facial expression while communicating.

Biocca and Harms [4] address the importance of nonverbal behavior in people's sense of access to each other's minds in mediated interaction situations. Nonverbal cues signal interactants' psychological and behavioral interdependence.

Clearly, the provision of visual nonverbal cues is an important element in producing communicator's copresence. But the question whether presenting such cues on the small screen of a mobile communication device is effective remains unanswered.

Presence, Copresence, and Social Copresence

There is a rich and sometimes contradictory literature defining the related concepts of "presence," "social presence," "copresence," and similar conceptualizations. It is well beyond the scope of this paper to review this literature in any depth, so we will instead mention a few typical conceptualizations before offering our own view of the important dimensions.

Biocca and Harms [4] assert that effective mediated interaction can be obtained by "the creation of social presence and the improvement of social communication." Biocca and his colleagues [5] further suggest the use of social presence measures to evaluate the performance of mobile systems, such as mobile video telephones.

The definition of social presence has evolved from a simple binary definition of feeling physically present or not to more of a continuum involving the "salience of others in mediated communication and consequent salience of their interpersonal interaction" [35]. It is important to consider interactions between human and mediated (or simulated) beings or environments as part of this definition. Biocca and his collaborators [5] provide a comprehensive definition of social presence, defining "social presence specifically to mean interactions in mediated environments" and focusing on social presence as a psychological concept. They elaborate the definition of social presence by adding copresence to the definition including "one's model of the other intelligence" and the sense of "access to another intelligence" that describes psychological engagement beyond the notion of physical awareness. This adds mutual understanding between interactants as an important component.

Swinth and Blascovich [39] state that the existing description of copresence has overlapped with the one of social presence [2,12]. Other researchers [24,34,39] differentiate the concept of social presence from the notion of copresence. Schroeder and his colleagues [34] further stress copresence as the concept which describes people's engagement in being and doing together in "a shared, computer-generated space" such as Collaborative Virtual Environments (CVEs) incorporating Virtual Reality (VR) technology. They differentiate this kind of mediated environment from ones displaying 2D graphics such as videophones or videoconferences. Swinth and Blascovich [39] provide a distinction between the concept of social presence and copresence found in communication studies and that found in the virtual environments. They point out that people can feel the sense of copresence when they are situated in physically different locations, but they perceive the others are "co-situated within the same interpersonal environment." Swinth and Blascovich [39] claim that actual social presence, which they view as having a physical dimension, is not required to feel copresence.

In our view, an expanded definition is needed that combines the concept of copresence championed by Biocca and his colleagues with the perceptions of the medium of communication by participants, and with an evaluation of the success of the communication by the participants. For want of a better term, we are calling this "Social Copresence." Social Copresence can be conceived of as involvement and engagement through mutual awareness between intelligent beings who have a sense of access to the other being consciously, psychologically and emotionally [4], within a mediated environment perceived as capable of supporting social communication.

Based on this general definition, we suggest that there are three elements that produce Social Copresence in the context of mediated communication. These are *Copresence*, defined as users' sense of being connected with their interaction partners [26]; *Social Richness of Medium*, defined as the perceived capability of the medium to connect interactants socially; and *Interactant Satisfaction*, defined as the presence of social attraction and emotional credibility between interactants.

Emotional credibility in particular has been missing from prior conceptualizations, which have focused more on interaction partners evaluating each other on the performance of pragmatic tasks. In this research, we will introduce measurement items that tap participants' affective evaluations of each other.

Anonymity and Avatars

Bailenson and his colleagues [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 to date. Some definitions include even static photo-realistic images of a person as an avatar (as we will, in the research reported here).

As a symbolic representation of a communicator, an avatar has two functions. The first is to represent the communicator and to communicate some information about him/her. In this case, a medium with a visual avatar is likely to produce more social copresence than a medium that does not include any visual representation of the communicator. The second function of an avatar is to mask the visual identity of the communicator, unless a photographic image or video of the communicator is used as an avatar.

Previous studies [2,11,12,26,27] have shown that the use of avatars promotes presence among users in some mediated environments. Satchell [33] has demonstrated a method to present users' social states using color coded avatars in mobile phone mediated conversation, for example.

The anonymity of communicating via an avatar has benefits. Milgram [22] states that "conditions of complete anonymity provide freedom from routinized social ties" and this represents the reason that urban dwellers maintain themselves anonymous from others to protect themselves from the dangerous social environment in cities. Turkle [41] has shown that people project alternate personae through anonymous virtual representations in mediated interactions. This identity construction via avatars is found in many mediated environments, including on-line games and other Interactive Virtual Environments (IVEs) and mobile social software applications for Mobile Phone Mediated Communication (MPMC). Other studies [2,3,11] have explored avatar uses in mediated interactions like group discussion and collaboration.

Avatars' nonverbal behavior provides a channel for transmitting the emotional signals [11] while not identifying the actual communicator. In their studies of human-computer communication with animated agents, some researchers [cf. 30, 42] have demonstrated that communication effectiveness is increased when people look at facial expressions during their interactions. Bailenson [2], for example, considers a videoconference image (unprocessed video) an avatar in some sense and claims that a videoconference is the most effective way to produce high realism in a mediated communication, presumably by providing a very realistic avatar that transmits facial expression. Avatar realism is generally defined as possession of either the visual appearance or behavioral attributes of a human [25,26,27]. Increased realism may enhance social copresence, but may also impair anonymity.

RESEARCH HYPOTHESES

There have been numerous studies [1,2,11,12,25,26,27] on the effects of avatar realism, both visual and behavioral. Most of these have involved collaborators or communicators using computer terminals to accomplish a task or achieve some pre-defined goal. The study reported here differs from that research in investigating the impact of different avatar presentations on a small screen typical of mobile communication devices, and in investigating the impact of avatars in a communication task that has an affective rather than an instrumental goal. It also focuses on investigating ways that the affective visual nonverbal component of communication can be preserved when the communicators wish to keep their visual identities private, as they might when conversing with strangers.

One fundamental way avatars may differ is in their behavioral realism, i.e., their actions. Both static (still image) and dynamic avatars have been used in mediated communication systems. Obviously, static avatars are less behaviorally realistic than dynamic ones. But behavioral realism can also vary in dynamic avatars. For example, facial expressions might be missing from a cartoon avatar. That would make it less behaviorally realistic, and less likely to effectively convey nonverbal cues relating to emotion. Kang, Watt and Ala [17] found that a communication system using dynamic avatars was perceived as more likely to provide the means of successful social interaction, although they did not find that dynamic avatars actually changed communication partners' affective evaluations of each other.

Dynamic avatars that convey facial expression would be expected to produce more social copresence than those that don't. Garau and her colleagues [12] report that avatars require high behavioral realism to increase users' sense of copresence. Other studies [1,2,11,26] demonstrate that communicators' social presence is reduced when they experience a large discrepancy between the behavioral realism and the visual realism of avatars, so more

behavioral realism should lead to higher social copresence. These findings lead to the following hypothesis:

- *H1: People will report higher Social Copresence (H1.1: Copresence; H1.2: Richness of Medium; and H1.3: Interactant Satisfaction) when interaction partners are represented by more behaviorally realistic dynamic avatars, rather than by static avatars.*

Another way avatars may differ is in their preservation of the identity of the communicator. Non-anonymous avatars are an accurate visual representation of the communicator. Anonymous avatars abstract, delete, or mask identity cues, and so would be expected to convey fewer nonverbal cues to a conversational partner. This leads to the following hypothesis:

- *H2: Interactants' Social Copresence (H2.1: Copresence; H2.2: Richness of Medium; and H2.3: Interactant Satisfaction) will be higher when interaction partners are represented by non-anonymous avatars, rather than by anonymous avatars.*

Anonymous avatars may also differ in their visual fidelity to the communicator they represent, or their *iconicity*. The most highly iconic avatar is a photographic representation of the communicator, but this would not be anonymous. A less iconic avatar might be a video or photo with degraded features to mask identity, while an even lower iconic avatar might be a cartoon caricature. Since lower iconic avatars would likely carry less nonverbal information about the communicator, we propose the following hypothesis:

- *H3: Interactants' Social Copresence (H3.1: Copresence; H3.2: Richness of Medium; and H3.3: Interactant Satisfaction) will be higher when interaction partners are represented by high-iconic anonymous avatars, rather than by low-iconic anonymous avatars.*

EXPERIMENTAL DESIGN

The basic experimental design was a 2 (behavioral realism) x 5 (visual realism or avatar type) x 3 (gender dyad type) factorial between-subjects experiment. The experimental conditions and the levels of each factor are shown in Table 1. To control for gender effects, three types of gender dyads were used in equal numbers in each experimental condition: Male-Male, Female-Female, and Male-Female. Differences in gender communication styles, while not the focus of this investigation, were explicitly included in the analysis to control for any potential confounding effects.

In the low behavioral realism conditions, subjects saw a static image displayed on a laptop computer showing a simulated mobile telephone (See Figure 1) with a realistic small-size screen size (~3.0") and display resolution (240 x 320 pixels). In the high behavioral realism conditions, the subjects saw video or dynamic animations on the same simulated screen. The laptop was placed at a normal arms-length distance from the subjects, simulating normal viewing conditions on an actual video mobile telephone. A commercial hands-free mobile telephone headset connected to the laptop was provided to users for the audio

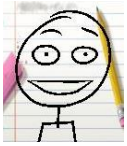
Behavioral Realism	Non-Anonymous Avatars	Anonymous Avatars			
		High-Iconic Avatars		Low-Iconic Avatars	
	Unmodified Video 	Overall Processed (edge-detector filtered) Video 	Outer Masked Processed Video 	High-Humanoid Avatar 	Low-Humanoid Avatar 
Dynamic Avatar (Video or Animation)	18 (3 gender dyads x 3 replications)	18 (3 gender dyads x 3 replications)	18 (3 gender dyads x 3 replications)	18 (3 gender dyads x 3 replications)	18 (3 gender dyads x 3 replications)
Static Avatar (Photo or Drawing)	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)	18 (3 gender dyads x 3 replications)

Table 1. Experimental conditions and numbers of participants.

conversation. The audio and video were combined within custom software that implemented the video processing required by each experimental condition, and transmitted the simulated video telephone signal over a Local Area Network connecting two isolated rooms in which the participants were located.

Examples of the avatars used in each experimental condition are shown in Table 1. The Overall Processed Video experimental condition was produced by implementing a real-time edge detector that processed the video for the simulated mobile telephone. Choice of this procedure was based on the findings of Zhao and Stasko [43]. Their results indicated that edge-detected video produced lower identification of video-recorded actors than the video pixelization that is often used to hide identities in news broadcasts [6]. Their results were confirmed in a pretest conducted as a preliminary to this research. In this pilot test, seventeen persons were asked to identify, from a set of sixteen photographs, the image of a person that they had viewed for several minutes in a video that had been edge-detection processed. Only 3 of 17 persons (18%) correctly identified the person, a result that was not statistically significantly higher than the percentage that would be expected to chose the correct photograph purely by chance (6%).

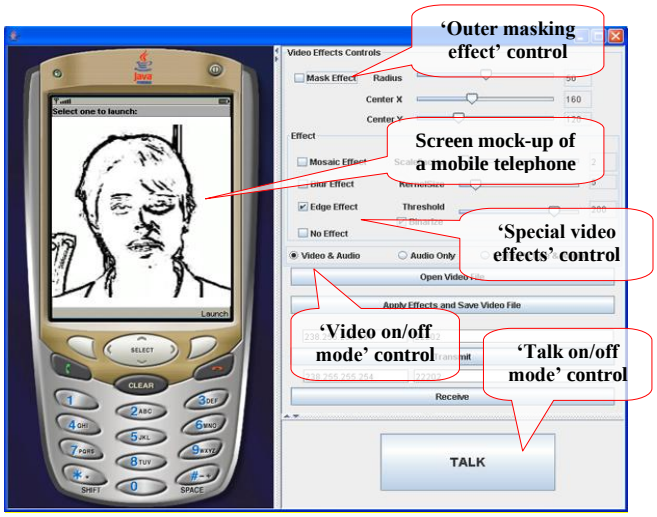


Figure 1. Mock-up of a mobile telephone and its control panel used to specify experimental conditions (control panel not visible during the experiment).

A second type of processed video based on the findings of Jarudi and Sinha [15] was used to produce the Outer Masked Processed Video experimental condition. Their research found that peripheral head cues, rather than central facial features, were most important in recognizing faces at low visual resolution. To give a condition of maximum anonymity, the edge-detected video was additionally masked to eliminate peripheral head cues in the Outer Masked Processed Video experimental condition.

To produce the animated high- or low-iconic avatar condition, a commercial product (Logitech QuickCam Orbit MP camera equipped with face-tracking function and "Logitech Video Effects" software) was used. This system generates a real-time animation of a graphical avatar's facial expressions that match the human communicator's expression, based on tracking and mapping, onto a graphic avatar 22 facial points from the webcam image of the human communicator. The avatar's facial expression then partially matches the human communicator's.

Experimental Participants and Procedures

One hundred eighty college students were recruited to participate in the experiment (see Table 1 for N's in each experimental condition). Each participant was paired with someone whom they were unlikely to know beforehand. The paired participants were randomly assigned to one of the experimental conditions. Participants were entered into a drawing for ten $20 gift certificates as incentive for participation. The experimental interaction took place in two separate rooms where the participants were placed separately and at different time, to eliminate initial face-to-face interaction.

Dyads were given a communication assignment that was deliberately chosen to be non-task oriented and to require self-disclosure and emotional discovery. Participants were given a hypothetical conversational scenario where they were instructed to assume the role of a student (not hard, since all were actually students) who seeks to find out if the other person is a suitable match with whom to share an apartment and possibly establish a friendly relationship. We propose that this conversation topic stimulates emotional interaction where people need to disclose personal and possibly sensitive information about themselves and to gather similar information about their conversational partner. Anonymity in this kind of self-revealing interaction with a stranger may be important, especially if the participants discover that they do not like each other. The conversational task lasted about ten minutes for the typical dyad.

Measurement Instruments[1]

Copresence. A seventeen item Likert-type scale was created with an 8-point metric for items (1 = Very Little; 8 = Very Much or 1 = Very Unlikely; 8 = Very Likely). All seventeen items were adopted from the items of copresence used in Nowak and Biocca's study [26]. This measurement included two separate sets of items: "perceived other's copresence (participants' perception of their interaction

[1] A full set of the measurement items is available from the corresponding author.

partners' involvement)" and "self-reported copresence (participants' self-report about their involvement)" [26]. Items included: "S/he was intensely involved in our interaction" and "I wanted to make the conversation more intimate."

Social Richness of Medium. A five item Likert-type scale was created with an 8-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 [26]. The items were worded to be applicable to a mobile telephone communication. Items included: "How likely is it that you would choose to use a mobile telephone interaction for a meeting in which you wanted to persuade the other person of something?" and "To what extent did you feel you could get to know someone that you met only through a mobile telephone?"

Interactant Satisfaction. A fifteen item Likert-type scale with an 8-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 [25]. Items included: "I would like to have a friendly chat with her/him" and "I think s/he could be a friend of mine." The other nine items, called emotional credibility, were newly created to measure interactants' emotional perceptions of their interaction partners (see [16,17,18]). The emotional perception items were derived from the evaluation dimensions of emotional intelligence test [37] which originated in the studies of Mayer and Salovey [32]. Items included: "S/he expresses feelings and emotions appropriately for the situation" and "S/he responds appropriately to positive and negative emotions."

RESULTS

Reliability tests showed good internal consistency among the 17 items of summed Copresence measure (Cronbach's alpha = .93), among the 5 items of the Perceived Social Richness of Medium measure (Cronbach's alpha = .86), as well as the 15 items of the Interactant Satisfaction measure (Cronbach's alpha = .90).

Between-subjects ANOVAs were conducted to test the hypotheses. Gender dyad type was included as a factor in all analyses, in order to discover and to control for any gender effects. Data for some subjects was not included, due to incomplete responses to Social Copresence measurement items.

H1: People will report higher Social Copresence (H1.1: Copresence; H1.2: Richness of Medium; and H1.3: Interactant Satisfaction) when interaction partners are represented by more behaviorally realistic dynamic avatars, rather than by static avatars.

H1.1: A three-way ANOVA (visual appearance by behavioral realism by gender dyad) results did not reveal a significant main effect ($p > .05$) on Copresence produced by behavioral realism (static vs. dynamic avatars). *Hypothesis H1.1 was not supported.*

However, there were statistically significant main effects for the visual appearance of the avatar [$F(4, 150)=5.15$, $p=.001$, eta-square=.12] and for the type of gender dyad [$F(2, 150)=3.12$, $p=.047$, eta-square=.04] on Copresence.

Post-hoc comparisons using the Tukey HSD test showed a statistically significant difference between the Low-Humanoid (low-iconic) anonymous avatar and the other avatars. The mean scores for the Unmodified Video ($M=5.75$, $SD=1.10$), the Overall Processed (edge-detector filtered) Video ($M=5.85$, $SD=.79$), Outer Masked Processed Video ($M=5.67$, $SD=.91$), and the High-Humanoid Avatar ($M=5.55$, $SD=1.10$) were not significantly different from each other, but all differed significantly from the Low-Humanoid Avatar ($M=4.92$, $SD=1.02$). Low-Iconic, Low-Humanoid Avatars produced significantly less Copresence between communicators.

Post-hoc comparisons using the Tukey HSD test showed that the mean score for the Female-Female dyads ($M=5.69$, $SD=.95$) and Male-Male dyads ($M=5.65$, $SD=1.01$) did not differ significantly, but both were higher than the Male-Female dyads ($M=5.29$, $SD=1.11$).

This indicates that participants were more satisfied with the conversation when interacting with the same gender partners when the effects of behavioral realism and visual appearance of the avatar were controlled. However, given that the low level of significance, and the lack of impact of gender dyads in most of the other hypothesis tests reported below, this finding must be interpreted with caution, as it may have been a statistical artifact.

There were no significant interaction effects (all F-values had $p > .05$).

H1.2: ANOVA results showed a large main effect for the avatar behavioral realism [$F(1,150)=23.43$, $p=.005$, eta-square=.14] on Social Richness of Medium. Dynamic avatars were judged significantly more capable of supporting social interaction ($M=4.98$, $SD=1.33$) than were Static avatars ($M=4.04$, $SD=1.27$). *H1.2 was supported.*

The results also showed a statistically significant main effect for the visual appearance of the avatar [$F(4, 150)=4.08$, $p=.004$, eta-square=.10] Post-hoc comparisons using the Tukey HSD test showed a statistically significant difference only for avatars' Iconic Type. The mean scores for the High-Humanoid Avatars ($M=4.08$, $SD=1.04$) and the Low-Humanoid Avatars ($M=4.07$, $SD=1.42$) were significantly lower than the Overall Processed (edge-detector filtered) Video ($M=5.03$, $SD=1.33$). Unprocessed Video ($M=4.84$, $SD=1.40$) and Outer Masked Processed Video ($M=4.53$, $SD=1.45$) did not differ significantly from any of the other avatar's visual appearance experimental conditions. The pattern of results indicates that a medium presenting more realistic avatars tended to be seen as more

capable of supporting social interaction, with abstract drawings being rated lower.

The results did not reveal a statistically significant main effect for the type of gender dyads on the measures of Social Richness of Medium (F-value had $p > .05$). There were no significant interaction effects (all F-values had $p > .05$).

H1.3: The results did not show a statistically significant main effect for realism of visual appearance ($p > .05$), behavioral realism (dynamic/animated or static), or the type of gender dyad on the measures of Interactant Satisfaction (all F-values had $p > .05$). There were no significant interaction effects ($p > .05$ for all F-values). *Hypothesis H1.3 was not supported.*

H2: Interactants' Social Copresence (H2.1: Copresence; H2.2: Richness of Medium; and H2.3: Interactant Satisfaction) will be higher when interaction partners are represented by non-anonymous avatars, rather than by anonymous avatars.

In this analysis non-anonymous avatars (unprocessed video or screen-capture photograph of the communicator) were contrasted with the other avatars that did provide anonymity.

H2.1: Three-way ANOVA results did not show a statistically significant main effect for avatar anonymity, behavioral realism (dynamic/animated or static), or the type of gender dyad on the measures of Copresence (all F-values had $p > .05$). No significant interaction effects emerged (all F-values had $p > .05$). *H2.1 was not supported.*

H2.2: The results did not reveal a statistically significant main effect for avatar anonymity or the type of gender dyad on the measure of Social Richness of Medium (all F-values had $p > .05$). No significant interaction effects (all $p > .05$) were found. *H2.2 was not supported.*

However, the results did show a statistically significant main effect for behavioral realism [$F(1, 168)=16.29$, $p<.001$, eta-square=.09] on this variable. Dynamic avatars were rated significantly higher ($M=4.98$, $SD=1.33$) than static avatars ($M=4.04$, $SD=1.27$).

H2.3: Three-way ANOVA results did not show a statistically significant main effect for avatar anonymity, behavioral realism (dynamic/animated or static), or the type of gender dyad on the measure of Interactant Satisfaction. There were no significant interaction effects (all F-values had $p > .05$). *H2.3 was not supported.*

H3: Interactants' Social Copresence (H3.1: Copresence; H3.2: Richness of Medium; and H3.3: Interactant Satisfaction) will be higher when interaction partners are represented by high-iconic anonymous avatars, rather than by low-iconic anonymous avatars.

In this analysis, the high-iconic anonymous avatar conditions (Overall Processed Video and Outer Masked Video) were contrasted with the low-iconic anonymous avatar conditions (High-Humanoid Avatar and Low-Humanoid Avatar) in a three-way ANOVA.

H3.1: The three-way ANOVA (avatar iconic type by behavioral realism by gender dyad) results revealed a statistically significant main effect for the iconic type of the avatar [$F(1, 132)=10.63$, $p=.001$, eta-square=.08]. High-iconic avatars produced significantly higher Copresence ($M=5.76$, $SD=.85$) than did Low-Iconic avatars ($M=5.23$, $SD=1.10$). *H3.1 was supported.*

The gender dyad main effect on Copresence was also significant [$F(2, 132)=4.16$, $p=.018$, eta-square=.06]. Post-hoc comparisons using the Tukey HSD test showed that the mean score for the Female-Female dyads ($M=5.75$, $SD=.95$) and Male-Male dyads ($M=5.55$, $SD=1.01$) were not significantly different, but both were significantly higher than the Male-Female dyads ($M=5.19$, $SD=1.03$). Again, this indicates that higher Copresence was reported in same-gender dyads.

The results did not reveal a statistically significant main effect for avatar behavioral realism on the measures of Copresence, nor were any interaction effects significant (all F-values had $p > .05$).

H3.2: The results showed a statistically significant main effect for the iconic type of the avatar [$F(1, 132)=11.02$, $p=.001$, eta-square=.08] on Social Richness of the Medium. Highly-iconic avatars were judged more capable of supporting social interaction ($M=4.78$, $SD=1.40$) than were low iconic avatars ($M=4.08$, $SD=1.24$). *H3.2 was supported.*

A main effect for behavioral realism was also found [$F(1, 132)=17.94$, $p=.000$, eta-square=.12] on Social Richness of Medium. Dynamic avatars were judged significantly higher ($M=4.88$, $SD=1.29$) on Social Richness of Medium than were static avatars ($M=3.98$, $SD=1.30$).

There were no statistically significant main effects for the type of gender dyad on the measure of Social Richness of Medium and no significant interaction effects were found (all F-values had $p > .05$).

H3.3: The results did not show a statistically significant main effect for the iconic type of the avatar, behavioral realism, or the type of gender dyad on the measure of Interactant Satisfaction. There were no significant interaction effects (all F-values had $p > .05$). *H3.3 was not supported.*

CONCLUSIONS AND DISCUSSION

The pattern of results of the hypothesis tests show the consistent impact of avatar visual and behavioral realism on participants' perceptions of the Social Richness of the Medium, i.e., the medium's ability to support social

interaction. Dynamic avatars were judged more capable of this support (H1.2 supported) as were High-Iconic avatars (H3.2 supported). Only the anonymity of the avatar made no difference in this perception (H2.2 not supported).

The impact of avatar differences extended somewhat less strongly to the actual personal impact of using the technology, however. There was no evidence that Dynamic avatars produced more Copresence than Static ones (H1.1 was not supported), or that Anonymous avatars produced less Copresence than Non-Anonymous ones (H2.1 was not supported). There was some effect of avatar visual differences, however, that was related to the visual realism of the avatar. High-Iconic avatars did produce a higher Copresence evaluation than did Low-Iconic avatars (H3.1 was supported), and video avatars (both unprocessed and also anonymized video) produced higher reports of Copresence than did graphical avatars. It appears that people feel more involved and get more engaged in their social interactions when avatars with more correct facial representation, expressions, and movement are used.

Most dramatically, avatar differences, be they behavioral realism, anonymity, or iconicity, showed no impact on participants' Interaction Satisfaction (H1.3, H2.3, and H3.3 were not supported). This element of Social Copresence does not appear to be sensitive to avatar characteristics, but is more likely produced by the actual content of the conversation, rather than the medium used to deliver the content.

The finding that people's perceptions of a medium diverge from their actual evaluations of the communication's success in a specific instance is consistent with the findings of Nowak, Watt and Walther [28], who found similar results in a task-oriented group collaboration experiment. This distinction between perception of the medium and evaluation of the actual communication is important, as some prior studies [31] have used perceptions of media appropriateness for social interaction as an indicator of social presence. The results of this study illustrate the incomplete nature of that evaluation process and make the case for including copresence and interaction satisfaction measures when evaluating the full Social Copresence potential of new technologies.

We speculated that people might prefer to use anonymous avatars to secure their privacy and be free from "social norms of politeness [11]," but that use of these avatars might impair the effectiveness of communication by masking or filtering nonverbal cues. The findings provide some reassurance that the process of introducing anonymity in the visual presentation does not degrade the Social Copresence of the communication. With other factors held constant, there was no difference seen between anonymous and non-anonymous avatars on any of the three elements of Social Copresence (H2.1, H2.2, and H2.3 were all not supported). In a related analysis, Kang, Watt, and Ala [17] also found no avatar type effect on interactant satisfaction with the communication. But they did find satisfaction was reduced when there was no visual provided, which would argue against simply turning off the visual transmission as an equally effective means of providing anonymity.

Combined with the finding that dynamic video avatars produce more Copresence in participants, the results imply that a video processing technique like the edge-detection filtering used in this experiment may be the best way to provide both anonymity and effective communication in mobile video telephones. This finding suggests that designers incorporate the option of an edge-detection anonymity filter into the interface of mobile telephones.

Previous studies [1,2,11,26] claim that users' social presence is lower when they experience a large incongruity between the behavioral realism and the visual realism of social actors (agents or avatars). Our findings tend to indirectly support this conclusion. Highly iconic (realistic) avatars that show normal movement (video as opposed to animation) produced higher evaluations on both Copresence and Social Richness of Medium.

Future Directions

This research adds to the growing body of knowledge about the conceptualization and measurement of social presence or copresence in mediated communication. The explicit measurement of emotional copresence adds a needed dimension to the full evaluation of the impact of a new communication technology. But extensions and refinements in the measurement of Social Copresence are clearly needed.

As discussed below, the results of this study, as well as results of most prior experimental studies, may be confounded by the type of communication task or setting used. We have no good taxonomy of communication or collaboration tasks within which to compare the results of different studies. The use of widely different communication tasks may be the root cause of some of the conflicting results of prior studies. Developing such a taxonomy, in which the relevant characteristics of each kind of communication task or setting are fully explicated, should be a high priority in future research.

One clear distinction in communication setting that was partially addressed in this research is the difference between a task-oriented communication and one where the goal is social information or cohesion. Mobile telephones in particular are a major medium for socially-oriented communication. The role of visual information is different in social communication than it is in accomplishing a specific task, but is important to both. Identifying the use and impact of visual information in the two contrasting communication settings is also a future research priority.

LIMITATIONS

This study examined communication between strangers in a single emotionally engaged situation. It is possible that these results are limited to that situation. It also involved role playing, and this may not produce responses that are identical to those produced in actual situations.

Some of the null results may have been induced by the kind of emotional and affective communication that was occurring. This type of communication may be less reliant on visual information, as opposed to the more pragmatic tasks addressed in prior studies. Alternatively, the lack of significant effects, especially when Copresence was being evaluated, may have been a function of the small visual display's limited abilities to convey nonverbal cues. However, while this may restrict the conclusions of this study to small-screen display devices, it is an accurate description of what occurs when using a mobile communication device, which will necessarily have a small display.

Using a simulation instead of an actual device means that the user's sensory experience was not identical to what the real device would deliver. For example, the eye gaze points might differ. However, the laptop seating arrangement for subjects placed the display close to the user's eyes, at about arm's length, where an actual device would likely be used, and the camera was placed at less than a 20-degree angle from the display, minimizing the differences from the actual device.

The experimental task was chosen to be relevant to college students, but use of this population still raises questions about the generalizability of the results to all users of mobile communication devices.

Previous studies [1,19] have documented users' gender differences in responding to avatar's eye gaze movement. The animation and facial expression tracking systems used in this study was not able to capture or reproduce eye gaze, so lack of this important nonverbal cue may have limited the observed effects of the low-iconic dynamic avatars in some gender combinations.

ACKNOWLEDGEMENTS

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

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. Bente, G. & Krämer, NC. Virtual gestures: Analyzing social presence effects of computer-mediated and computer-generated nonverbal behaviour. *Fifth Annual International Workshop PRESENCE 2002* (2002).
4. Biocca, F. & Harms, C. What is social presence? In F. Gouveia & F. Biocca (Eds.), *Proc. Presence 2002.* University of Fernando Pessoa Press (2002).
5. 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).
6. 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).
7. Cumiskey, K. Mobile fantasies on film: Fostering "co-presence" through the integrated use of real and imagined mobile technology. *Pre-Conference, International Communication Association* (2006).
8. Eagle, N. & Pentland, A. Social serendipity: Mobilizing social software. *IEEE Pervasive Computing,* 28-34, April-June 2005.
9. Ekman, P. & Friesen, W. Nonverbal Behavior in psychotherapy research. *Research in psychotherapy: proceedings of the third conference*, vol 3: 179-183 (1968).
10. Ekman, P. & Friesen, W. The repertoire of nonverbal behavior: Categories, origins, and coding. *Semiotica*,1: 49-98 (1969).
11. Garau, M. *The Impact of Avatar Fidelity on Social Interaction in Virtual Environments*. Ph.D. thesis, Department of Computer Science, University College London (2003).
12. 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).
13. Humphreys, L. Mobile Devices and Social Networking. *Pre-Conference, International Communication Association,* (2006).
14. Izard, C. E. Emotions and facial expressions: A perspective from differential emotions theory. In Russell, J. A. & Fernandez-Dols, J. M., *The psychology of facial expression* (pp. 57-77). Cambridge University Press (1997).
15. Jarudi, I. & Sinha, P. Relative contributions of internal and external features to face recognition (2003). Available at http://cbcl.mit.edu/cbcl/publications/ai-publications/2003/AIM-2003-004.pdf.

16. Kang, S. The impact of digital iconic realism on anonymous interactants' mobile phone communication. *Proc. CHI'07.* ACM Press (2007).
17. Kang, S., Watt, J., & Ala, S. Communicators' perceptions of social presence as a function of avatar realism in small display mobile communication device. *Proc. Hawaii International Conference on System Sciences* (2008).
18. 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).
19. Knapp, M. & Hall, J. *Nonverbal communication in human interaction*. Harcourt Brace Jovanovich College Publishers (1992).
20. Ling, R. *The mobile connection: The cell phone's impact on society*. Morgan Kaufmann Publishers (2004).
21. Littlejohn, S. *Theories of Human Communication*. Wadsworth (2002).
22. Milgram, S. *The Individual in a Social World: Essays and Experiments*. McGraw Hill, Inc. (1977).
23. Ministry of Social Development in New Zealand. Social Connectedness. *Social Report 2006* (2006). Available at http://www.socialreport.msd.govt.nz/social-connectedness/index.html.
24. Nowak, K. Defining and differentiating copresence, social presence and presence as transportation. Paper presented at the *Presence 2001 Conference* (2001). Available at http://citeseer.ist.psu.edu/cache/papers/cs/26451/http:zSzzSznimbus.temple.eduzSz~mlombardzSzP2001zSzNowak1.pdf/nowak01defining.pdf.
25. Nowak, K. L. The Influence of anthropomorphism and agency on social judgment in virtual environments. *Journal of Computer Mediated Communication, 9*(2), (2004). Available at http://jcmc.indiana.edu/vol9/issue2/nowak.html.
26. 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).
27. 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). Available at http://jcmc.indiana.edu/vol11/issue1/nowak.html.
28. Nowak, K. L., Watt, J., & Walther, J. B. The influence of synchrony and sensory modality on the person perception process in computer-mediated groups. *Journal of Computer-Mediated Communication, 10*(3), article 3 (2005). Available at http://jcmc.indiana.edu/vol10/issue3/nowak.html.
29. Nyiri, K. *Mobile Democracy: Essays on society, self and politics*. Vienna: Passagen Verlag (2003).
30. Poggi, I. & Pelachaud, C. Performative facial expressions in animated faces. In Cassell, J., Sullivan, J., Prevost, S. & Churchill, E. (eds.) *Embodied conversational agents*. The MIT Press (2000).
31. Rice, R. E. Media appropriateness: Using social presence theory to compare traditional and new organizational media. *Human Communication Research, 19*, 451-484. (1993).
32. Salovey, P., Brackett, M., & Mayer, J. *Emotional Intelligence: Key readings on the Mayer and Salovey Model*. Dude Publishing (2004).
33. Satchell, C. "Contextualising Mobile Presence with Digital Images" In PICS Workshop, *Ubicomp* (2006).
34. Schroeder, R., Steed, A., Axelsson, A-S., Heldal, I., Abelin, A°. Widestro¨m, J., Nilsson, A., & Slater, M. Collaborating in networked immersive spaces: As good as being there together? *Computers & Graphics*, 25:781–788, (2001).
35. Short, J., Williams, E., & Christie, B. *The social psychology of telecommunication*. London: John Wiley & Sons (1976).
36. Smith, I. E. Social mobile applications. *IEEE Computer Society*, 38(4):84-85, April 2005.
37. Smith, J. Emotional Intelligence Report: EIQ16 Questionnaire (2004). Available at http://www.myskillsprofile.com/eiq16_sample_report.pdf.
38. Smith, E. & Mackie D. *Social Psychology*, 2nd Edition. Psychology Press, New York, USA (2000).
39. Swinth, K. & Blascovich, J. Perceving and responding to others: Human-human and human-computer social interaction in collaborative virtual environments. In *Fifth Annual International Workshop: Presence 2002*: p. 310–340 (2002).
40. Tomita, H. Geitai and the Intimate Stranger. In Ito, M., Okabe, D. & Matsuda M. (eds.). *Personal, Portable, Pedestrian*. The MIT Press (2005).
41. Turkel, S. *Life on the screen: Identity in the age of the Internet*. Touchstone (1995).
42. 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).
43. Zhao, Q. & Stasko, J. Evaluating image filtering based techniques in media space applications. *Proc. the 1998 ACM conference on Computer supported cooperative work* (1998).