Jeremy N. Bailenson
Department of Communication
Stanford University
Stanford CA 94305 USA
Bailenson@stanford.edu

Kim Swinth
Department of Psychology
University of California, Santa Barbara
Santa Barbara CA 93106 USA

Crystal Hoyt
School of Leadership
University of Richmond
Richmond VA 23173 USA

Susan Persky
Alex Dimov
Jim Blascovich
Department of Psychology
University of California, Santa Barbara
Santa Barbara CA 93106 USA

Presence, *Vol. 14, No. 4, August 2005, 379–393*
© 2005 by the Massachusetts Institute of Technology

The Independent and Interactive Effects of Embodied-Agent Appearance and Behavior on Self-Report, Cognitive, and Behavioral Markers of Copresence in Immersive Virtual Environments

Abstract

The current study examined how assessments of copresence in an immersive virtual environment are influenced by variations in how much an embodied agent resembles a human being in appearance and behavior. We measured the extent to which virtual representations were both perceived and treated as if they were human via self-report, behavioral, and cognitive dependent measures. Distinctive patterns of findings emerged with respect to the behavior and appearance of embodied agents depending on the definition and operationalization of copresence. Independent and interactive effects for appearance and behavior were found suggesting that assessing the impact of behavioral realism on copresence without taking into account the appearance of the embodied agent (and vice versa) can lead to misleading conclusions. Consistent with the results of previous research, copresence was lowest when there was a large mismatch between the appearance and behavioral realism of an embodied agent.

1 Introduction

The use of *avatars* (i.e., digital representations of human users) to represent individuals within multiuser virtual environments in real time has increased considerably in the past decade, underscoring the need to better understand how users experience mediated interactions with others. In addition, the use of *embodied agents* (i.e., digital representations of computer programs that have been designed to interact with, or on behalf of, a human) as interactive guides or as interactants within virtual environments has prompted inquiry into how people perceive and respond to nonhuman social entities.

In this paper, we examine how definitions and operationalizations of *copresence* (also commonly referred to as social presence) are influenced by variations in the human resemblance of embodied agents in terms of appearance and behavior within an immersive virtual environment. In doing so, we explore how and when definitional and measurement differences regarding copresence lead to convergent and divergent findings with respect to the independent and interactive effects of the appearance and behavior of embodied agents. Our multimethod approach included self-report, behavioral, and cognitive dependent measures.

1.1 Defining and Operationalizing Copresence

One of the difficulties in understanding and synthesizing the accumulating literature on copresence is the plethora of definitions and operalizations used by investigators, as can be seen in the thorough discussion by Schroeder (2002). Unfortunately, this difficulty has led to inconsistent and often contradictory findings reported in the copresence literature. Heeter (1992), for example, defined copresence (or social presence) as the extent to which other beings, both living and synthetic, exist in a virtual world and appear to react to human interactants. Nowak (2001), on the other hand, defined copresence as the sense of connection with another mind. Russo (2001) argued that copresence (or what she called *mediated presence*) is the perception by a communicator that another person in a mediated or online environment is real, immediate, or present. Conversely, Slater and his colleagues have defined copresence as the sense of being and acting with others in a virtual place (Slater, Sadagic, Usoh, & Schroeder, 2000). Finally, Blascovich and his colleagues have defined copresence as the extent to which individuals treat embodied agents as if they were other real human beings (Blascovich et al., 2002).

Although there is no consensual definition of copresence, or agreement about how best to assess it, existing research on copresence can be roughly divided into two camps: studies that focus on people's *perceptions* of embodied agents versus studies that focus on people's *social responses* to embodied agents. Researchers from the first camp (e.g., Heeter, 1992; Biocca, 1997; Lombard & Ditton, 1997; Garau, Slater, Bee, & Sasse, 2001; Lessiter, Freeman, Keogh, & Davidoff, 2001 define copresence in terms of a user's perception or feeling that others are cosituated within an interpersonal environment (Swinth & Blascovich, 2002). Researchers in this tradition take the view that copresence occurs when a user not only perceives or feels that he or she is "present" within a virtual environment, but that he or she perceives or feels that others are present there as well. Because this view of copresence focuses on people's phenomenological experiences, researchers adopting this orientation frequently operationalize copresence in terms of a user's self-reported perception and/or feeling that they are not alone in a virtual environment.

Another group of researchers, however, views copresence as a social, task-related, or physiological *response* to embodied agents (e.g., Bailenson et al., 2004; Blascovich et al., 2002; Reeves & Nass, 1996; Bailenson, Beall, & Blascovich, 2002; Beall, Bailenson, Loomis, Blascovich, & Rex, 2003; Blascovich et al., 2002; Lee & Nass, 2004; Bente, Rüggenberg, Tietz, & Wortberg, 2004; Beall, Bailenson, Loomis, Blascovich, & Rex, 2003; Meehan, Insko, Whitton, & Brooks, 2004; Slater, 2004; Bente, Rüggenberg, Tietz, & Wortberg, 2004). Researchers adopting this orientation argue that copresence occurs when people treat embodied agents as if they were other real people, and they operationalize copresence using a variety of methods including the analysis of physiological data, examination of a variety of nonverbal and verbal behaviors, or the assessment of task performance, cooperation, interaction satisfaction, liking, or other social responses.

In the current study, participants were asked to examine embodied agents that varied in behavior and appearance. We then measured the extent to which the embodied agents were perceived and treated as if they were human via self-report, behavioral responses, and cognitive dependent measures. Adopting this approach enabled us to compare the two general ways of defining copresence described above by examining both participants' self-reported *perceptions* of embodied agents as well as their self-reported, cognitive, and behavioral *responses* to them. Moreover, by employing three distinct measurement modalities (self-report questionnaires, behavioral responses, and cognitive responses) we were able to distinguish converging and diverging patterns pertinent to each of the two general ways of defining copresence. We discuss each of these response modalities in turn.

1.2 Self-Report Markers of Copresence

In order to tap the phenomenological aspects of copresence, we administered a self-report questionnaire that asked participants to report on the extent to which they perceived or felt like there was another social entity with them in the virtual environment. In addition, we administered two self-report measures designed to assess different kinds of social responses people might have to embodied agents: liking of an embodied agent and willingness to perform embarrassing acts in front of an embodied agent. We hypothesized that all three indicators of copresence would be correlated, positively for copresence and liking, and negatively between copresence/liking and willingness to perform embarrassing acts.

1.3 Behavioral Markers of Copresence

In order to assess the idea that one can infer copresence by observing interaction with an embodied agent, we assessed a behavioral measure of copresence: interpersonal distance. Building on previous work (e.g., Reeves & Nass, 1996; Krikorian, Lee, Chock, & Harms, 2000; Bailenson, Blascovich, Beall, & Loomis, 2001; Bailenson, Blascovich, Beall, & Loomis, 2003), we hypothesized that copresence would be positively related to hesitation in approaching an embodied agent, interpersonal distance, and maintenance of interpersonal distance.

1.4 Cognitive Markers of Copresence

Furthermore, we assessed memory for embodied agents. Previous research has found that people remember more features about people in scenes than about nonsentient objects in scenes (New, 2003). Consequently, users should remember more information about embodied agents represented as sentient beings (i.e., human) than about nonsentient types of representations. Thus, we expected better memory for embodied-agent features. However, Hoyt, Blascovich, and Swinth (2003) demonstrated that the presence of an embodied agent worsened performance on a cognitive task when participants performed in front of an audience of avatars, and that perceived copresence mediated that effect. Consequently, it could be argued that copresence may, in fact, be indicated by worse memory for information relating to an embodied agent. In the current study, we examined users' memory for an embodied-agent feature as a cognitive marker of copresence on an exploratory basis, without predictions concerning the direction of memory effects. Our goal in including this cognitive marker of copresence was to compare any memory differences observed to other, more established and predictable markers of copresence in order to begin to understand the cognitive effects of copresence, if any.

1.5 Factors Affecting Copresence

Although a number of factors may influence copresence, two, appearance and behavior of embodied agents, have received considerable attention. Although some work has examined the interplay between how human an embodied agent appears and how an embodied agent behaves (e.g., Bailenson et al., 2001; Garau et al., 2003), researchers typically examine each factor in isolation or confound them as a single variable.

Bailenson et al. (2001), for example, found that making an embodied agent more realistic looking by adding a texture map from an actual photograph of a face did not increase copresence compared with an embodied agent that was more cartoonlike, as long as both types of embodied agents demonstrated realistic gaze behaviors. Similarly, Garau et al. (2003) demonstrated that more photographically realistic avatars elicited no more copresence than did less realistic-looking avatars, and, in fact, found that increasing the photo-realism of an avatar can actually cause a decrease in copresence if the behavioral realism is not also increased correspondingly (see also Slater & Steed, 2001; Garau, 2003).

In contrast, Nowak and her colleagues (Nowak, 2001; Nowak & Biocca, 2001) demonstrated that the extent to which embodied agents resemble real human beings affects people's social judgments of agents as interaction partners and the amount of copresence experienced. In one study, interaction partners represented by anthropomorphic images were rated as more socially attractive and they received higher partner satisfaction ratings than did partners represented either by no image or by an image that did not appear human (Nowak, 2001). In addition, participants indicated that they felt greater copresence and that the communication medium was better able to support the social interaction when interactants were represented by highly anthropomorphic images compared to when no image or an image low in anthropomorphism was used (Nowak & Biocca, 2001). Similarly, Parise, Kiesler, Sproull, & Waters (1996) demonstrated that more human-looking embodied agents elicited more liking behavior and more cooperation than did less human-looking embodied agents.

One possible explanation for these seemingly contradictory findings concerns previous manipulations of appearance. Studies that have demonstrated null effects have typically manipulated the visual fidelity (or photo-realism) of embodied agents and studies that have found a significant effect of embodied-agent appearance on copresence and other social outcomes have typically manipulated the anthropomorphism of the embodied agent, or the extent to which the embodied agent appears human in form. We are unaware of studies that have systematically varied both photo-realism and anthropomorphism and, unfortunately, most studies that have been conducted on the effect of embodied-agent appearance have tended to confound photo-realism and anthropomorphism.

Given that studies focusing on the photo-realism of an embodied agent have consistently reported null effects, however, we attempted to control for the effect of photo-realism in the current study by using embodied agents of equivalent visual fidelity while simultaneously manipulating how much like an actual human being the embodied agent both looked and behaved. We predicted that embodied agents that either looked human or acted human would elicit greater copresence than embodied agents that did not look or act human. However, based on the work of Blascovich and his colleagues (Blascovich, 2002; Blascovich, et al., 2002), we also hypothesized that embodied-agent appearance and behavior would interact with each other such that the behavioral realism of an embodied agent would matter most for non-human-looking embodied agents.

1.6 Summary and Study Objectives

In this study, we systematically examined the independent and interactive effects of both the appearance and behavior of embodied agents on users' perceptions and social responses within an immersive virtual environment. In doing so, we attempted to address a number of currently unresolved issues. First, we examined how different ways of defining and operationalizing copresence are influenced by systematic manipulations of appearance and behavior of virtual representations. In doing so, we attempted to reconcile some of the contrasting findings that have been reported in the research literature. Second, we examined whether similar or disparate patterns of response obtained across different modalities of copresence measurement (i.e., self-report vs cognitive vs behavioral dependent measures). Finally, we used four different algorithms to control the head movements of the embodied agent, some of which have never before been empirically compared in an immersive virtual environment.

2 Method

2.1 Participants

One hundred and forty-six undergraduates (73 males, 73 females) ranging in age from 18 to 27 years (M = 19.63 years) recruited from the University of California, Santa Barbara, served as participants. They were either paid or received class credit for participation.

2.2 Design

A two-way between-subjects factorial design was used. The first factor, *representation type*, included three levels of embodiment, human, teddy bear, and blockhead. As can be seen in Figure 1, the three embodied agents varied in the extent to which they resembled a

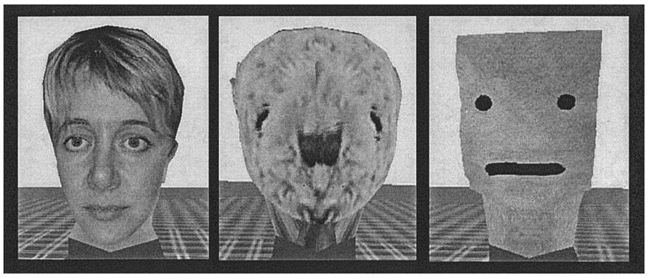

Figure 1. *Photographically realistic human, teddy bear, and blockhead embodied agents.*

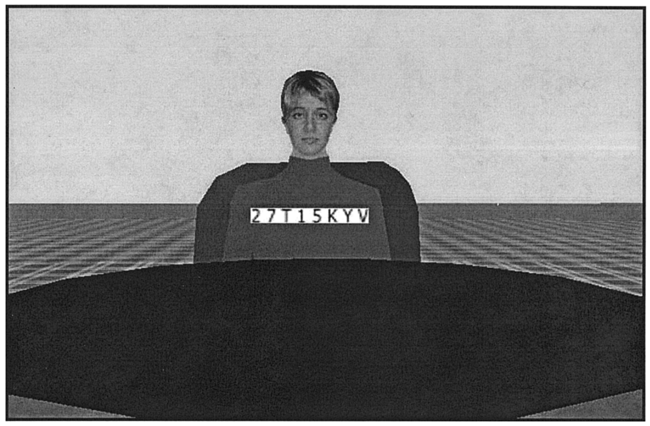

Figure 2. *A participant's view of the virtual room.*

human being.[1] For female participants, the human representation created using 3dMeNow software, depicted in the left panel of Figure 1, was used. For male participants, a similar-looking male representation was used. All three representations possessed equivalent behavioral functionality within the virtual environment (see below) and exhibited similar levels of photo-realism.

While we attempted to ensure as much as possible equal levels of photo-realism (by using textures of similar resolution and by ensuring that the underlying 3D mesh for each model contained similar numbers of anchor points and structural facial features), subjectively there may have been minor differences in photo-realism.

The second between-subjects factor, *behavioral realism,* included four levels of head movements. Four different algorithms were employed to control the embodied agent's head movements. In the *static* condition, the embodied agent remained motionless. In the *random movement* condition, the embodied agent's head continually moved in trajectories generated by the computer toward random points within a 60-degree arc of each side of its head (for both pitch and yaw). In the *mimic* condition, the embodied agent's head movements mirrored the participant's actual head movements with a 4-s delay.[2] In the *recorded movement* condition, head movements recorded from a randomly selected participant in a previous session were used to control the embodied agent's head movements. In all four conditions, the embodied agent blinked its eyes according to a preset algorithm but there were no facial gestures exhibited. Complete crossing of these two independent variables resulted in 12 between-subjects conditions.[3]

2.2 The Virtual Environment

The virtual room in which participants were immersed is depicted in Figure 2. Participants sat in a chair, at a table, facing an embodied agent. Participants were able to see the head and shoulders of the agent. The same body was used for each representation type. A tag containing a string of letters and numbers, later used as part of a surprise memory task, was positioned on the shirt of the embodied agent.

1. There are a variety of terms that researchers use to describe how real or how human a virtual representation looks, including anthropomorphism, visual realism, and photographic (or photo-) realism. We have chosen to use the more general term *representation type* as a variable name and talk about the manipulation as mainly pertaining to photographic realism. We acknowledge that other terms, such as anthropomorphism, however, may also be appropriate.

2. Previous research using mimicked head movements has indicated that users generally do not detect embodied agents that are mimicking their head movements with a 4-second delay (Bailenson, Beall, Loomis, Blascovich, & Turk, 2004; Bailenson & Yee, in press). However, this study is unique in that it directly compared the effect of mimicked head movements with other head-movement algorithms.

3. There were an unequal number of females and males across the 12 conditions.

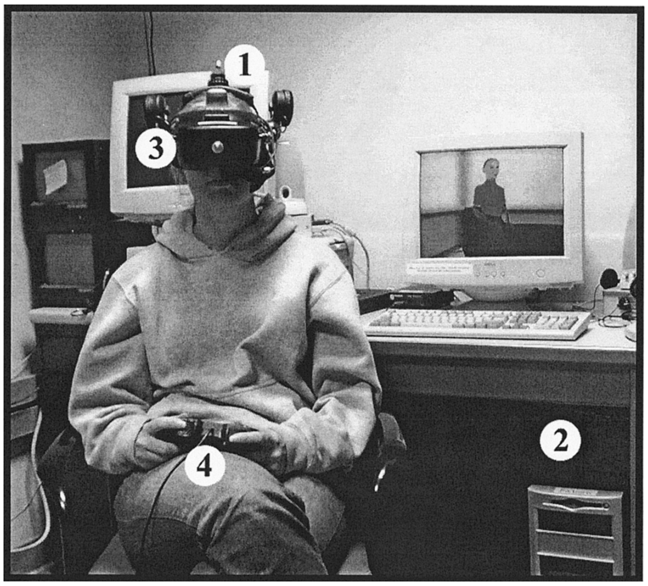

Figure 3. *A depiction of our immersive virtual environment system. The components are: (1) orientation tracking sensor, (2) image generator, (3) HMD, and (4) game pad input device.*

2.3 Equipment

Participants wore a Virtual Research V8 stereoscopic head-mounted display (HMD; see Figure 3). The HMD featured dual 680 horizontal by 480 vertical pixel resolution LCD panels that refreshed at 72 Hz. The display optics presented a visual field subtending approximately 50 degrees horizontally by 38 degrees vertically.

Perspectively correct stereoscopic images were rendered by a 450 MHz Pentium III dual-processor computer with an Evans & Sutherland Tornado 3000 dual pipe graphics card, and were updated at an average frame rate of 36 Hz. The simulated viewpoint was continually updated as a function of the participants' head movements. The orientation of the participant's head was tracked by a three-axis orientation sensing system (Intersense IS300, update rate of 150 Hz). The system latency, or the amount of delay between a participant's head movement and the resulting concomitant update in the HMD's visual display, was 65 ms maximum. The software used to assimilate the rendering and tracking was Vizard 2.0. Participants used a Logitech Rumble-Pad Pro game pad to interact with the virtual environment.

2.4 Procedure

When participants arrived at the laboratory, an experimenter sat them down and instructed them on how to wear and adjust the HMD and how to use the game pad depicted in Figure 3 to interact with the virtual environment. Once immersed, participants found themselves seated at a table directly across from an embodied agent (see Figure 2). Depending on the assigned condition, the embodied agent varied in appearance and behavior. Using the game pad, participants scrolled through the following instructional text that appeared above the embodied agent's head:

> In this laboratory, we are interested in people's reactions to "others" in virtual environments. In this study, you will spend time with an "other." After a short time, we will ask you questions about the "other." You will use the game pad to answer the questions. The questions measure your agreement. The more you agree, the higher the number you should choose on the rating scale in front of you.

We carefully chose and repeatedly used the word "other" set off by quotation marks to avoid using suggestive nouns such as "person" or "agent." Participants were next instructed to look directly at the "other" for 90 s. Next, a series of 10 *copresence, embarrassment,* and *likability* self-report questionnaire items (see Appendix A) appeared one at a time over the embodied agent's head in random order. Participants indicated their agreement with these items by using the game pad to select a response on a Likert-type rating scale that appeared under each item. Hence, ratings were provided in real time as participants experienced the embodied agent.

After completing the items, participants were instructed to approach the embodied agent for further examination. They used buttons on the game pad to move toward (positive translation on the z axis) or away from (negative translation on the z axis) the embodied agent, with each press of a game-pad button equivalent to a jump of 2.5 cm (the starting distance between the participant's nose and the embodied agent's nose was approximately 3 m). We did not render the body of participants, so there was no leaning or walking animation during the translation—only their viewpoint moved closer or farther from the agent. After participants finished approaching the embodied agent and examining it, they took off the HMD and filled out a demographic questionnaire. Finally, we administered a surprise memory test in which participants were asked to recall the random string of characters written on the embodied agent's chest (see Figure 2).

2.5 Measures

2.5.1 Perceived Copresence. Perceived copresence was assessed via a three-item self-report measure. We created a composite copresence measure ($\alpha = .71$; see Appendix A) by averaging the three copresence questions. Higher scores indicated higher levels of self-reported copresence where ratings were made on a Likert-type scale with response options ranging from 0 (*strongly disagree*) to 6 (*strongly agree*).

2.5.2 Embarrassment. Embarrassment was assessed as a self-reported social response to the presence of an embodied agent. Previous research has demonstrated that willingness to perform embarrassing acts in front of a virtual representation is a worthy measure of that representation's degree of social influence (Bailenson et al., 2003). We averaged the three embarrassment questions ($\alpha = .72$; see Appendix A) into a single, composite measure, with higher numbers indicating greater willingness to perform embarrassing acts in front of an embodied agent.

2.5.3 Likability. Likability was assessed as a second self-reported social response to the presence of an embodied agent. We averaged the four likability questions ($\alpha = .71$; see Appendix A) into a single, composite measure, with higher numbers indicating that participants liked the virtual representation more.

2.5.4 Memory. Memory was quantified as the total number of correct characters (out of a possible eight) a participant recalled from the virtual representation's label, without penalty for incorrect answers.

2.5.5 Interpersonal Distance. Interpersonal distance was included as a behavioral social response to the presence of an embodied agent. We assessed two indicators of interpersonal distance: *minimum distance* and *reversal count*. Minimum distance was the closest distance the participant came to the embodied agent; reversal count referred to the number of times a participant changed direction along the z-axis (i.e., went back and forth) during their examination of the embodied agent. We consider reversal count to be a proxy for hesitation in approaching the embodied agent.

3 Results and Discussion

3.1 Hypotheses

We hypothesized that greater copresence would be indicated by higher perceived copresence ratings, greater liking of the embodied agent, less willingness to commit embarrassing acts in front of the embodied agent, the maintenance of greater interpersonal distance, and more hesitation in approaching the embodied agent. We did not, however, have any specific predictions with respect to the relationship between copresence and memory.

Based on the results from previous copresence studies, we expected greater copresence (across our various copresence measures) to occur with realistically behaving embodied agents (specifically our recorded and mimic behavioral-realism conditions that featured movements based on actual human motion) compared with unrealistically behaving embodied agents (i.e., our static and random-movement conditions). In addition, we expected greater copresence to occur with more human-looking embodied agents than with less human-looking ones. Drawing on predictions derived from the Blascovich et al. (2002) threshold model of social influence, however, we expected embodied-agent appearance to moderate the effect of behavioral realism such that behavioral realism would matter more for less human-looking embodied agents than for more human-looking embodied agents. Finally, we expected to demonstrate convergence in our dependent measures such that they would correlate with one another and change in the same way as a function of our independent variable manipulations.

3.2 Perceived Copresence Ratings

Results of a 3 (representation type: human vs. teddy bear vs. blockhead) × 4 (behavioral realism: static vs. random movement vs. mimic vs. recorded) analysis of covariance, controlling for participant gender,[4] revealed a significant main effect of behavioral realism on self-reported perceived copresence, $F(3, 133) = 2.72$, $p < .05$, $\eta^2 = .06$. As Figure 4 demonstrates, perceived copresence was higher in the random-movements condition compared to the static condition, $p < .01$. However, this main effect of behavioral realism was qualified by a significant representation type by behavioral-realism interaction, $F(6, 133) = 2.18$, $p < .05$, $\eta^2 = .09$.

Simple main effects tests revealed that within the blockhead condition, perceived copresence varied significantly as a function of behavioral realism, $F(3, 133) = 3.39$, $p < .05$, $\eta^2 = .07$. Specifically, perceived copresence was significantly higher when the blockhead representation moved according to our random-movement algorithm compared with either the static or mimic conditions, *p*s < .05. Simple main effects tests also revealed that perceived copresence varied as a function of behavioral realism for participants in the teddy-bear condition, $F(3, 133) = 3.30$, $p < .05$, $\eta^2 = .07$. Specifically, perceived copresence was significantly higher in mimic condition than in either the static or recorded conditions, *p*s < .05, and higher in the random-movements condition relative to the recorded and static conditions, (*p*s < .05 and <.10 respectively).

In contrast, there were no significant differences in perceived copresence between any of the four behavioral-realism conditions for participants in the human "other" condition. This pattern of findings is consistent with our prediction that behavioral realism would have less of an influence on perceived copresence for embodied agents that appear human, as compared to embodied agents with a nonhuman appearance. In addition, it is interesting to note that the type of head movement that elicited the greatest level of perceived copresence varied as a function of the embodied agent's appearance, with participants attributing high copresence ratings to both mimicking and randomly moving teddy-bear representations but only randomly moving blockhead representations.

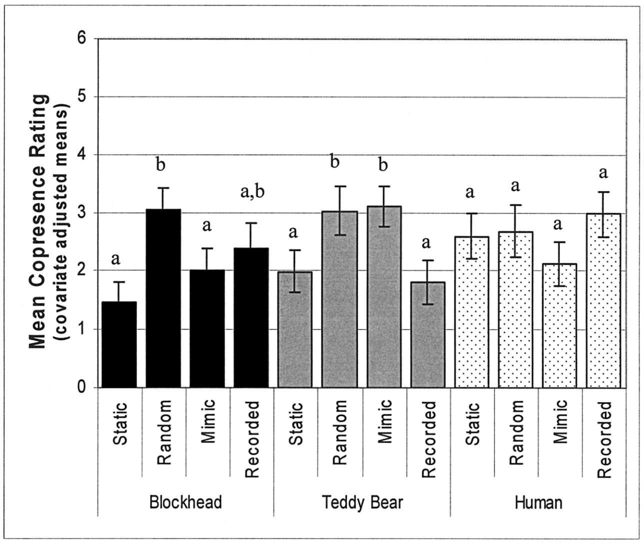

Figure 4. *Mean perceived-copresence ratings (using gender-covariate adjusted means) as a function of representation type and behavioral realism. Conditions that share the same letter do not differ within that level of representation type at p < .05.*

4. Although both male and female participants were used in this study, gender was not included as a separate factor in analyses involving both behavioral realism and representation type. When randomly assigning participants to conditions, we did not block participants by gender and, as a result, when behavioral realism and representation type were fully crossed, there were not enough participants of both genders in all 12 experimental conditions to reliably compute a three-way analysis of variance. Thus, our data analytic technique involved including participant gender as a covariate in our analyses involving both behavioral realism and representation type in order to control for the potential effect of participant gender. The results do not change appreciably if we remove gender as a covariate in our analyses, but we include gender as a covariate to provide the most thorough analysis possible.

3.3 Embarrassment Ratings

Results of a 3 (representation type: human vs. teddy bear vs. blockhead) $\times$ 4 (behavioral realism: static vs. random movements vs. mimic vs. recorded) analysis of covariance, controlling for participant gender, revealed a significant main effect of representation type on participants' self-reported willingness to perform embarrassing acts in front of the embodied agent (i.e., embarrassment), $F(2, 133) = 4.21$, $p < .05$, $\eta^2 = .06$. Specifically, participants in the human embodied-agent condition ($M = 2.60$, $SD = 1.55$) were significantly less willing to perform embarrassing acts in front of the "other" than were participants in the teddy-bear ($M = 3.51$, $SD = 1.54$) embodied-agent condition, $p < .05$. Participants in the human-representation condition were also marginally less willing to perform embarrassing acts than participants in the blockhead ($M = 3.13$, $SD = 1.56$) condition, $p < .10$. In sum, the presence of a human-looking representation seemed to reduce participants' willingness to perform in embarrassing acts. Regardless of the behavioral realism of the representation, embodied agents who looked like humans elicited greater social influence than did non-human-looking representations.

3.4 Likability Ratings

Results of a 3 (representation type: human vs. teddy bear vs. blockhead) $\times$ 4 (behavioral realism: static vs. random movements vs. mimic vs. recorded) analysis of covariance, controlling for participant gender, also revealed a significant main effect of representation type on self-reported likability ratings, $F(2, 133) = 5.43$, $p < .01$, $\eta^2 = .08$. Specifically, both the human-looking representation ($M = 2.50$, $SD = 1.09$) and the teddy-bear representation ($M = 2.28$, $SD = 1.09$) were rated as significantly more likable than the blockhead representation ($M = 1.79$, $SD = 1.10$), ps $< .05$. Participants found the blockhead representation (which was the least human looking of the three representations) to be less attractive than either of the other two representation types and they liked the blockhead representation the least.

3.5 Correlations between Perceived Copresence, Embarrassment, and Likability Ratings

In order to examine the relationships among perceived copresence, embarrassment, and likability, a series of Pearson correlations was computed. Consistent with out predictions, results of these analyses revealed a significant negative correlation between perceived copresence and a participant's willingness to perform embarrassing acts in front of the embodied agent, $r(144) = -.27$, $p < .001$ as well as a significant positive correlation between perceived copresence and likability, $r(144) = .31$, $p < .001$. Specifically, we found that the higher the perceived copresence experienced, the less willing participants were to perform embarrassing acts in front of the embodied agent ($r^2 = .07$) and the more participants liked the embodied agent ($r^2 = .10$). There was no significant correlation between likability and embarrassment ratings.

3.6 Memory

We computed a 3 (representation type: human vs. teddy bear vs. blockhead) $\times$ 4 (behavioral realism: static vs. random movements vs. mimic vs. recorded) analysis of covariance, controlling for participant gender, to examine the effects of our independent variables on participants' memory. Results of these analyses revealed a significant main effect of behavioral realism on memory, $F(3, 133) = 4.14$, $p < .01$, $\eta^2 = .09$. Specifically, memory was significantly worse in the mimic condition than in either the static or random-movement conditions, ps $< .05$, suggesting that, overall, participants had the worst memory for characters worn by a representation that mimicked their own movements with a 4-s delay. Again, however, this main effect was qualified by a significant representation type by behavioral-realism interaction (see Figure 5), $F(6, 133) = 2.70$, $p < .02$, $\eta^2 = .11$.

Simple main effects tests revealed that within the blockhead condition, memory varied significantly as a function of behavioral realism, $F(3, 133) = 2.71$, $p < .05$, $\eta^2 = .06$. Specifically, individuals remembered significantly more characters when they were in the

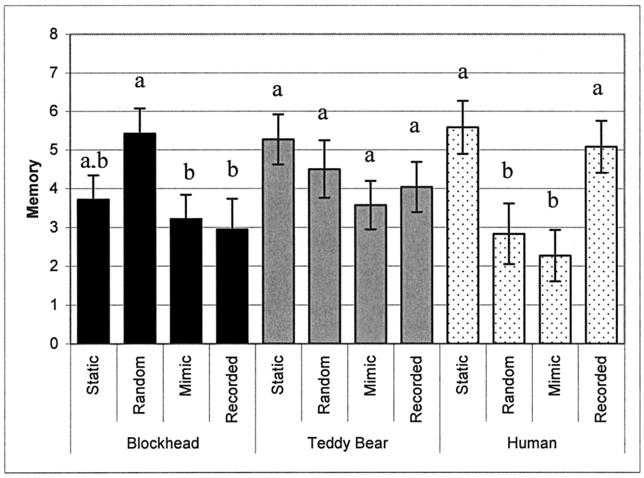

Figure 5. *Gender-covariate adjusted mean number of correctly recalled characters out of eight on a surprise memory test as a function of representation type and behavioral realism. Conditions that share the same letter do not differ within that level of representation type at p < .05.*

random-movement condition than in either the mimic or the recorded conditions (the two behavioral-realism conditions that appeared most humanlike), *p*s $< .05$, and participants in the random-movement condition remembered marginally more than participants in the static condition, $p < .10$. These results converge quite well with the perceived copresence data, in which participants indicated higher copresence ratings for blockhead representations only when the blockhead demonstrated random head movements. Interestingly, this suggests that for representations who look the least human, the least human-looking head movements were seen as most appropriate. This finding, that avoiding mismatches is the most effective manner of combining an embodied agent's appearance and behavior, is consistent with previous research on visually embodied agents (Slater & Steed, 2001; Garau et al., 2003) as well as voice agents (Lee & Nass, 2003).

Furthermore, the convergence of memory and perceived-copresence ratings provides some evidence that better memory is indicative of higher copresence. In other words, the more copresence an embodied agent elicits, the more a person interacting with that embodied agent remembers about it. It is interesting to note, however, that despite the similar patterns of findings between the perceived-copresence data and the memory data, results of a Pearson correlation revealed that there was no significant linear relationship between perceived copresence and people's ability to remember information about an embodied agent.

There was, however, a significant simple main effect of behavioral realism within the human-representation condition, $F(3, 133) = 5.49$, $p < .001$, $\eta^2 = .11$. Specifically, participants in the mimic and random-movement conditions remembered significantly fewer characters than participants in either the static or recorded conditions, *p*s $< .05$. Participants in the teddy-bear condition displayed the same pattern, but in that condition, none of the differences reached statistical significance.

These results suggest that participants may have been distracted by a human-looking embodied agent that mimicked their own movements, which, in turn, affected their ability to notice and recall information in the virtual environment. Moreover, participants may also have been distracted in the random-movement condition since this condition represented the greatest mismatch between the appearance (human looking) and behavior (not human looking) of the embodied agent. Further research should examine these potential explanations as they suggest that different processes may be involved for different algorithms for generating embodied-agent head movements.

3.7 Interpersonal Distance

There were no significant effects of behavioral realism or representation type on minimum distance (our first measure of interpersonal distance). Thus, participants in all conditions maintained equivalent levels of interpersonal distance between themselves and the embodied agent. There was also no significant correlation between perceived copresence and either of our behavioral dependent measures (i.e., minimum distance and number of returns).

However, as hypothesized, results of a 3 (representation type: human vs. teddy bear vs. blockhead) $\times$ 4 (be-

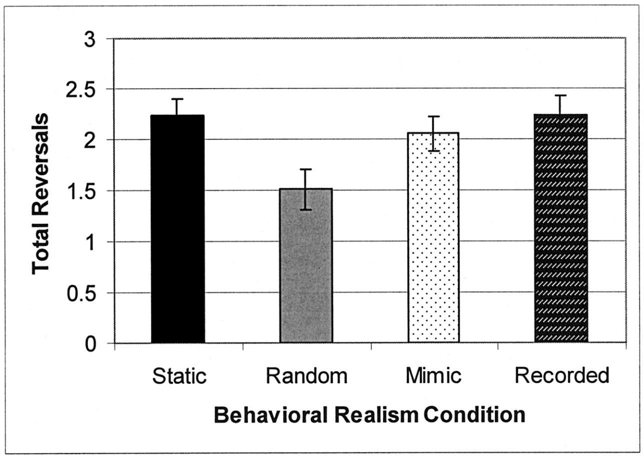

Figure 6. *Mean number of direction reversals while approaching the embodied agent as a function of behavioral realism.*

havioral realism: static vs. random movements vs. mimic vs. recorded) analysis of covariance, controlling for participant gender, revealed a significant main effect of behavioral realism on reversal count, or the number of times a participant changed direction along the *z*-axis during his or her examination of the embodied agent, $F(3, 133) = 3.06$, $p < .05$, $\eta^2 = .07$. As can be seen in Figure 6, participants in the random-movement condition exhibited significantly fewer reversals than participants in any of the other three behavioral-realism conditions. Pairwise comparisons indicated that there were significantly fewer reversals in the random-movement condition compared to the other three conditions ($ps < .05$), which are not significantly different from each other.

This finding, in part, confirms our hypothesis concerning our behavioral dependent variable. For an embodied agent whose head moves with curves randomly generated by the computer, participants show little to no hesitation in approaching. In contrast, when the head movements were controlled by a real person (either a mimic of their own movements or the movements from another participant), participants showed greater hesitation in approaching the embodied agent. Thus, computer-generated movements generated less hesitancy than human-generated movements.

Given this explanation, however, one might expect even less hesitation, that is, fewer reversals, in approaching the static representation. This finding did not occur, and one potential explanation is that, overall, a randomly moving representation may elicit less social influence and copresence than a completely static representation. In other words, the static representation has at least the potential to exhibit realistic movements, while the unrealistically moving embodied agent immediately lowers the sense of copresence. In sum, participants' hesitancy in approaching an embodied agent was influenced by the behavioral realism of the embodied agent, but not its appearance.

4 Conclusions

4.1 Summary of Findings

To summarize, the appearance and behavior of embodied agents had varying effects on the dependent variables. In many instances, there was convergence, in that self-report, cognitive, and behavioral measures changed in similar ways as a function of our independent variable manipulations. For example, the perceived-copresence, likability, and embarrassment measures correlated in the predicted directions, and the blockhead representation elicited similar reactions across behavior types for both the memory measure and the perceived-copresence measure. In other instances, however, there was a lack of convergence where varying realism produced different outcomes for the various response modes. One inevitable challenge facing researchers who examine a large number of dependent variables is to understand the instances in which those dependent variables do and do not converge.

In the current study, for some of our dependent measures of copresence (willingness to commit embarrassing acts and likability), manipulation of representation type affected copresence responses. On other dependent measures (self-reported copresence ratings, interpersonal distance approach reversals, and memory), manipulation of behavioral realism changed participants' responses. One way to interpret these results is that self-report measures were adequate for measuring copresence levels as affected by variation in appearance whereas self-report measures were not sensitive enough to detect the subtle differences elicited by the behavioral-realism manipulation. Indeed, the only reliable main effects of the behavioral-realism manipulation occurred with cognitive and behavioral dependent measures.

In other words, self-report questionnaires are effective when measuring how people *perceive* an embodied agent, but not necessarily how they will *respond* to that embodied agent. Consequently, not only does it matter how one defines copresence, it also matters how one operationalizes copresence. If the goal of a researcher is to understand the way in which people fathom embodied agents on a conceptual level, then self-report questionnaires may be the best tool. However, if the goal of the researcher is to understand the way in which people will interact with, learn from, or behave toward an agent, then he or she should augment questionnaires with other types of measures. In this paper, we are not lobbying for the abandonment of copresence self-report questionnaires; however, we do urge that researchers use these questionnaires as a complement to other types of measures whenever possible.

While this explanation that self-report measures lack sensitivity receives support from the current data, it must be qualified. On most measures, there was an interaction between behavioral realism and representation type. At times, these interactions did not conform to our specific predictions and were difficult to interpret. Part of the reason for this difficulty is most likely due to the fact that we assumed a priori that certain types of head movements would be viewed as more realistic than others. Specifically, we assumed that the measures based on actual human behaviors (mimic and recorded) would appear more realistic than either those generated by the computer (random) or than no movements at all. Recorded human movements applied to 3D models, however, may not always be the most effective method of rendering realistic behaviors, and the differences between our various levels of behavioral realism may have been too subtle in the current study. As previous work has demonstrated (Garau et al., 2003), the interaction between appearance and behavioral realism is complicated, and the finding that seems to be emerging is that large disparities between the two types of realism can cause diminished levels of copresence. Only when the blockhead representation, the least humanlike of the three, demonstrated unrealistic, random head movements did it elicit high levels of self-report and cognitive copresence.

Our findings with respect to the main effects of appearance on self-report measures suggest that how human an embodied agent looks may also be an important precondition for experiencing copresence. This finding is unanticipated given predictions made by the Blascovich model and given findings from previous work (Bailenson et al., 2001; Garau et al, 2003). It seems that virtual representations that look like a realistic human, as opposed to either a familiar or unfamiliar nonhuman, elicit embarrassment reactions that would be appropriate only in the presence of a real human.

4.2 Limitations and Future Research

Though this effect surfaces for representations that are humanlike to varying degrees, how this effect of appearance would apply to interactions with multiple human representations of varying photographic realism remains an empirical question. One key direction for future research is to examine how variation in photographic realism (i.e., visual fidelity) of humanlike embodied agents affects copresence and how it interacts with behavioral realism. Furthermore, another important direction is to begin future studies with pretested levels of all realism variables (behavioral, photographic, and others) to assist in a more thorough interpretation of results. Moreover, the current study is limited in that we implemented behavioral realism solely by varying head movements. A more thorough investigation of this phenomenon should investigate other types of behaviors, such as facial gestures, hand movements, and gait, as well as more complex behaviors such as planning and coordinated actions. In addition, the interaction task we chose (a purely nonverbal task) is limited, and research should definitely examine copresence in richer social-interaction settings. Finally, our method of translation, using the game pad as opposed to naturally tracked walking, was not a perfect behavioral measure, since people do not translate by pushing buttons in the real world. Future work should provide a further examination of copresence using more natural behavioral measures.

Importantly, had we measured only self-report responses, we would have demonstrated a different pattern of results in the current study. Attempting to augment any study geared toward understanding human interaction with embodied agents or avatars with dependent measures from multiple response channels certainly can make data interpretation more complicated. Patterns from self-report data that seem clean and consistent often are not replicated when a less direct measurement paradigm is implemented. Furthermore, when asking peripheral, implicitly related questions about presence, such as likability or embarrassment, one risks measuring constructs inside the human mind that are not "central" to the definition of copresence. Adding more implicit and disparate measures may not paint a cleaner picture, but it does paint a more complete one. The challenge to the researcher is interpreting these complex patterns in the data in order to assess a more thorough understanding of an extremely novel and abstract phenomenon of interacting with a virtual, digital entity.

It may be the case that copresence is a purely affective construct inside the human mind; alternatively, copresence may be more of a cognitive construct. In this early stage of the research field examining copresence in virtual environments, the theoretical and empirical work is extremely limited, and does not allow one to rule out either of these two (or many other) possibilities. Determining the exact conceptual nature of an abstract construct in the human mind is a challenge that faces many social scientists. The measurement framework we offer allows the researcher to get a more complete picture of the copresence construct, and allows for future research to utilize more specific types of construct-type hypotheses as well as measurement strategies.

In addition to the basic research implications with respect to the use of various modes of dependent measures, the current study also has some important applied implications. By understanding the components of a virtual representation that lead to elicitations of heightened copresence with that representation, we can better develop our virtual agents and virtual worlds to immerse and engage the user. As we, as well as others, have demonstrated, creating high levels of copresence is a complicated process involving an appropriate fit between levels of each type of realism. As we grow to better understand these interactions, we can better match agent design to context and perhaps reach a level where it becomes easy to use agents to develop the degree of copresence usually found with virtual avatars, and perhaps, in the long term, the degree of copresence found with flesh-and-blood people during face-to-face interaction.

References

Bailenson, J. N., Aharoni, E., Beall, A. C., Guadagno, R., Dimov, A., & Blascovich, J. (2004). *Comparing behavioral and self-report measures of embodied agents social presence in immersive virtual environments.* Proceedings of the 7th Annual International Workshop on Presence, Valencia, Spain.

Bailenson, J. N., Beall, A. C., & Blascovich, J. (2002). Mutual gaze and task performance in shared virtual environments. *Journal of Visualization and Computer Animation, 13,* 1–8.

Bailenson, J. N., Beall, A. C., Loomis, J., Blascovich, J., & Turk, M. (2004). Transformed social interaction: Decoupling representation from behavior and form in collaborative virtual environments. *Presence: Teleoperators and Virtual Environments, 13*(4), 428–441.

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

Bailenson, J. N., Blascovich, J., Beall, A. C., & Loomis, J. M. (2003). Interpersonal distance in immersive virtual environments. *Personality and Social Psychology Bulletin, 29,* 1– 15.

Bailenson, J. N., & Yee, N. (in press). Digital chameleons: Automatic assimilation of nonverbal gestures in immersive virtual environments. *Psychological Science.*

Beall, A. C., Bailenson, J. N., Loomis, J., Blascovich, J., & Rex, C. (2003). Non-zero-sum mutual gaze in collaborative virtual environments. *Proceedings of HCI International,* Crete.

Bente, G., Rüggenberg, S., Tietz, B., & Wortberg, S. (2004). *Measuring behavioral correlates of social presence in virtual encounters.* Paper presented at the International Communication Association Conference, May 27–31.

Biocca, F. (1997). Cyborg's dilemma: Progressive embodiment in virtual environments. *Journal of Computer-Mediated Communication, 3*(2). Available from: http://www.ascusc.org/jcmc/vol3/issue2/biocca2.html.

Blascovich, J. (2002). Social influences within immersive virtual environments. In R. Schroeder (Ed.), *The social life of avatars*. London: Springer-Verlag.

Blascovich, J., Loomis, J., Beall, A., Swinth, K., Hoyt, C., & Bailenson, J. N. (2002). Immersive virtual environment technology as a methodological tool for social psychology. *Psychological Inquiry, 13,* 103–124.

Garau, M. (2003). *The impact of avatar fidelity on social interaction in virtual environments*. Doctoral thesis, University College London, London, England.

Garau, M., Slater, M., Bee, S., & Sasse, M. A. (2001). The impact of eye gaze on communication using humanoid avatars. *Proceedings of the SIG-CHI Conference on Human Factors in Computing Systems,* Seattle, WA.

Garau, M., Slater, M., Vinayagamoorthy, V., Brogni, A., Steed, A., Sasse, M. A. (2003). The impact of avatar realism and eye gaze control on perceived quality of communication in a shared immersive virtual environment. *Proceedings of the SIG-CHI Conference on Human Factors in Computing Systems,* Fort Lauderdale, FL.

Heeter, C. (1992). Being there: The subjective experience of presence. *Presence: Teleoperators and Virtual Environments, 1*(2), 262–271.

Hoyt, C. L., Blascovich, J., & Swinth, K. R. (2003). Social inhibition in immersive virtual environments. *Presence: Teleoperators and Virtual Environments, 12*(2), 183–195.

Krikorian, D. H., Lee, J., Chock, T. M., & Harms, C. (2000). Isn't that spatial? Distance and communication in a 2-D virtual environment. *Journal of Computer-Mediated Communication, 5*(4). Available from: http://www.ascusc.org/jcmc/vol5/issue4/krikorian.html.

Lee, K. M., & Nass, C. (2003). Designing social presence of social actors in human computer interaction. *Proceedings of the CHI 2003 Conference on Human Factors in Computing Systems,* 289–296.

Lee, K. M., & Nass, C. (2004). The multiple source effect and synthesized speech: Doubly disembodied language as a conceptual framework. *Human Communication Research, 30,* 182–207.

Lessiter, J., Freeman, J., Keogh, E., & Davidoff, J. D. (2001). A cross-media presence questionnaire: The ITC sense of presence inventory. *Presence: Teleoperators and Virtual Environments, 10*(3), 282–297.

Lombard, M., & Ditton, T. (1997). At the heart of it all: The concept of presence. *Journal of Computer-Mediated Communication, 3*(2). Available from: http://www.ascusc.org/jcmc/vol3/issue2/.

Meehan, M., Insko, B., Whitton, M., Brooks, F. P. (2002). Physiological measures of presence in stressful virtual environments. *ACM Transactions on Graphics, 21*(3), 645–653.

New, J. J. (2003). A content-specific attenuation of change blindness: Preferential attention to animate beings in natural scenes. *Journal of Vision, 3*(9), 643. Available from: http://journalofvision.org/3/9/643/.

Nowak, K. (2001, May). *The influence of anthropomorphism on social judgment in social virtual environments*. Paper presented at the International Communication Association, Washington, DC.

Nowak, K., & Biocca, F. (2001, May). *The influence of agency and the virtual body on presence, social presence, and copresence in computer mediated interaction*. Paper presented at the 4th International Workshop on Presence, Philadelphia, PA.

Parise, S., Kiesler, S. B., Sproull, S., & Waters, K. (1996). My partner is a real dog: Cooperation with social agents. *Proceedings of Computer Supported Cooperative Work,* 399–408.

Reeves, B., & Nass, C. (1996). *The media equation: How people treat computers, television, and new media like real people and places*. New York: Cambridge University Press.

Russo, T. C. (2001, May). *Operationalizing mediated presence. Initial steps toward a measure of the construct*. Paper presented at the 4th International Workshop on Presence, Philadelphia, PA.

Schroeder, R. (2002). Social interaction in virtual environments: Key issues, common themes, and a framework for research. In R. Schroeder (Ed.), *The social life of avatars: Presence and interaction in shared virtual environments,* (pp. 1–18). London: Springer-Verlag.

Slater, M. (2004). How colorful was your day? Why questionnaires cannot assess presence in virtual environments. *Presence: Teleoperators and Virtual Environments, 13*(4), 484–493.

Slater, M., Sadagic, A., Usoh, M., Schroeder, R. (2000). Small group behavior in a virtual and real environment: A comparative study. *Presence: Teleoperators and Virtual Environments, 9,* 37–51.

Slater, M., & Steed, A. (2001). Meeting people virtually: Experiments in virtual environments. In R. Schroeder (Ed.), *The social life of avatars: Presence and interaction in shared virtual environments* (pp. 146–171). London: Springer-Verlag.

Swinth, K. R., & Blascovich, J. (2002). Perceiving and responding to others: Human-human and human-computer social interaction in collaborative virtual environments. *Pro-*

ceedings of the 5th Annual International Workshop on PRESENCE, Porto, Portugal.

Appendix: Questionnaires

Note: (r) denotes reverse-coded items

***Copresence* (α = .71)**

1. Even when the "other" was present, I still felt alone in the virtual room (r).
2. I felt like there was someone else in the room with me.
3. I felt like the "other" was aware of my presence in the room.

***Embarrassment* (α = .72)**

1. I would be willing to change clothes in front of the "other."
2. I would be willing to pick my nose in front of the "other."
3. I would be willing to act out a scene from the movie "Titanic" in front of the "other."

***Likability* (α = .71)**

1. I like the "other."
2. I would like to meet this "other" again.
3. The "other" is attractive.
4. Spending time with the "other" was NOT satisfying (r).