Perceiving and Responding to Others: Human-Human and Human-Computer Social Interaction in Collaborative Virtual Environments

Kimberly R. Swinth & Jim Blascovich
Research Center for Virtual Environments and Behavior
Department of Psychology
University of California
Santa Barbara, CA 93106
swinth@psych.ucsb.edu

Keywords: social presence, copresence, social verification, social interaction, mediated communication, computer-mediated communication, technology-mediated interaction, agents, embodied agents, avatars

Abstract

In this paper, we propose a framework for understanding technology-mediated social interaction that attempts to synthesize theory and research from the fields of social psychology, communications, sociology, and computer science. Although this perspective has developed primarily to describe the social dynamics of interactions occurring within collaborative virtual environments (CVEs), we believe this general framework will be of use to those who study other types of mediated and non-mediated interactions as well.

1. Introduction

The use of multi-user, shared, or collaborative virtual environments (CVE) for education, training, entertainment, and collaboration among geographically dispersed work groups and individuals has expanded considerably in the past several years, bringing with it a need to understand more thoroughly how users of this as well as other telecommunications technologies experience their mediated interactions with others. In addition, the use of computerized agents (i.e., a computer program that has been designed to interact with, or on behalf of, a human) as interactive website guides or as social interactants within virtual environments, along with the expanded use of avatars to represents human users within CVEs, has prompted inquiry into how people perceive and respond to non-human social entities (e.g., agents) and to virtually-represented others (e.g., avatars).

In this paper, we address two issues that we believe have important implications for successful CVE research, design, and implementation. First, we review the concepts of social presence and copresence as they are most frequently used in the
communications/media studies and virtual environment/computer science literatures. In doing so, we note that the lack of a consensual definition and the confounding of the constructs makes it difficult to integrate research findings both within and across disciplines or to develop valid and reliable measures of them. Building on research and theorizing in the fields of social psychology and sociology, we then propose working definitions of social presence and copresence that attempt to integrate the multiple perspectives and which helps clarify the relationships and distinctions between the two constructs.

Next, we propose a general theoretical model of social influence that attempts to explain when and how the presence of others leads to various forms of social influence within both technology-mediated and non-mediated social interactions. We note that although the mere presence of others is sometimes sufficient to induce social influence, mere presence alone is an insufficient explanation for the full range of social influence phenomena that have been documented. Thus, we argue that when studying how the presence of others affects people’s thoughts, feelings, and behaviors, it is important to not only examine the extent to which one perceives that others are co-situated within some interpersonal environment, but to also identify factors that mediate and moderate the effect of the presence of others on various social psychological outcomes.

Finally, we address the issue of how people perceive non-human social entities, such as computerized agents, as well as how they respond to virtually-represented (or embodied) others within CVEs. The idea that humans respond socially to agents and computers has been the topic of a considerable amount of recent research attention (Allbeck & Badler, 2002; Bailenson, Blascovich, Beall, & Loomis, 2001; Bailenson, Blascovich, Beall, & Loomis, in press; Blascovich, 2002a; 2002b; Blascovich, Loomis, Beall, Swinth, Hoyt, & Bailenson, 2002; Brent & Thompson, 1999; Clark, 1999; Hoyt, Blascovich, & Swinth, in press; Moon & Nass, 1996; Nass & Moon, 2000; Nass, Moon, Fogg, Reeves, & Dryer, 1995; Nass, Steuer, & Tauber, 1994; Nowak, 2001; Nowak & Biocca, 2001; Pertaub, Slater, & Barker, 2002; Reeves & Nass, 1996). However, there is some disagreement about how users experience these kinds of human-human and human-computer interactions. Some researchers argue that humans’ social responses to computers and agents are just like their social responses to other real people (e.g., Moon & Nass, 1996; Nass & Moon, 2000; Nass, et al., 1995; Nass et al., 1994; Nowak, 2001; Nowak & Biocca, 2001; Reeves & Nass, 1996). Others, however, have found that this is not always the case (e.g., Bailenson et al., 2001; Bailenson et al., in press;
We attempt to reconcile these two sets of seemingly contradictory findings by discussing factors that we believe moderate people’s perceptions of and social responses to computerized-agents and humans embodied by avatars.

2. Social Presence in the Communications Literature

One issue that communications theorists and researchers have grappled with over the past several decades is how and how much technological mediation affects individuals’ perceptions of and responses to one another. The concept of social presence was first proposed by communications researchers to describe the extent to which telecommunications media differ in their ability to convey the salience of others in a mediated interpersonal interaction and the consequent salience of their interpersonal relationship (Short, Williams, & Christy, 1976). According to this perspective, telecommunications media differ in their ability to transmit social information and social cues, and social presence reflects the degree to which a medium is perceived as conveying the presence of others. The extent to which the social presence of the interactants is conveyed depends not only on the words exchanged during the interaction, but also on a range of both nonverbal and verbal cues and the interactional context. Communications researchers frequently use social presence to compare face-to-face and mediated interactions or mediated interactions to one another (e.g., Short et al., 1976, Walther, 1992).

According to Short and his colleagues (Short et al., 1976), telecommunications media differ in their capacity to transmit information about proximity, orientation, the physical appearance and attractiveness of interactants, facial expressions, direction of looking, mutual eye-gaze, posture, dress, and various other nonverbal and verbal cues. These objective differences in media quality contribute to subjective differences in the social presence elicited by those media. Moreover, they suggest that the fewer channels of information available within a given medium, the less attention users pay to the presence of other social participants. Thus, Short and his colleagues argue that social presence is an important key to understanding person-to-person telecommunications in that it varies between different media, it affects the nature of the interaction, and it interacts with the purpose of the interaction to influence the medium chosen by an individual who wishes to communicate.
Similarly, Sproull and Kiesler (1986) argue that the critical difference between face-to-face and mediated communications is that many forms of media lack social context cues. According to these researchers, social context cues include aspects of the physical environment as well as actors' nonverbal behaviors that define the nature of the social situation and the actors' relative status within it. In face-to-face settings, for example, these cues might be conveyed by spatial features, artifacts, and physical adornments (see Burgoon, Buller, & Woodall, 1989 for a thorough review of nonverbal aspects related to situation and status cues). Moreover, Kiesler, Siegel, and McGuire (1984) argue that it is the absence of such cues in many types of computer-mediated communication (CMC) that leads to uninhibited communication such as "flaming" (i.e., insults, swearing, and hostile, intense language); greater self-absorption versus other-orientation; and messages reflecting status equalization or equalized participation (e.g., when actors, who might otherwise defer speaking turns to higher-status participants, become disinhibited and speak out of turn).

Like social presence theory (Short et al., 1976) and social context cues theory (Sproull & Kiesler, 1986), media richness theory (Daft & Lengel, 1984, 1986; Daft, Lengel, & Trevino, 1987; Trevino, Daft, & Lengel, 1990; Trevino, Lengel, & Daft, 1987) argues that communication characteristics differ across various media based on the bandwidth or the number of cue systems available within them. Face-to-face interaction is touted as the "richest" form of communication, given the availability of immediate feedback, the number of cues and channels utilized, nonverbal (facial and oral) backchanneling cues, personalization, and language variety. On the other hand, computer-mediated communications, such as text-based electronic bulletin boards and e-mail, are considered to be very "lean", because nonverbal cues are not present (however, see Godin, 1993 for an alternative view). Other media, such as video and audio recordings, are described as moderately rich according to their channel capacities.

An underlying principle in all three of these theoretic traditions is that a good match between media characteristics and the goal of the interaction will lead to "better" (more effective, satisfying, etc.) communication. Moreover, all three theories suggest that media differ in the extent to which they are appropriate for different types of interactions with some media offering very limited utility while others are appropriate in a greater number of contexts (see also Rice, 1993).
However, despite the theoretical convergence among these three approaches, empirical research has produced somewhat contradictory results regarding the role of social presence in technology-mediated interaction. As predicted by social presence (Short et al., 1976), social context cues (Sproull & Kiesler, 1986), and media richness (Daft & Lengel, 1984; 1986) theories, some studies have found that technology-mediated communications are less personal or socioemotional than face-to-face interactions (Hiltz, Johnson, & Turoff, 1986; Connolly, Jessup, & Valacich, 1990; Rice & Love, 1987) while others have not found this to be the case (e.g., Kerr & Hiltz, 1982; Hiltz & Turoff, 1978; Walther, 1992, 1997). In addition, some studies (e.g., Tammelin, 1998) have failed to find differences in social presence ratings across media.

One reason for these contradictory findings may be that although it is true that media differ in the kinds and amount of social information they transmit and that this may affect the salience of social interactants to one another, these theories tend to largely overlook the fact there are likely to be a number of variables that moderate peoples’ experience with a given communicative episode. For example, the relationship between interactants, their interaction history with one another, their beliefs about one another, their interaction goals, and the extent to which they each perceive that the other is attentive and responsive may all affect users’ perceptions and responses, regardless of the medium used to interact. In addition, there are likely to be a variety of motivational, affective, personality, and demographic characteristics that also affect a user’s attention, reactions, and motivation to interact.

Although research into these types of moderators is limited, there is some support for the idea that peoples’ perceptions of one another and their feelings about using a given medium to interact may change with experience and time. For example, Walther and his colleagues (Walther, 1992; 1997; Walther & Burgoon, 1992; Walther & Tidwell, 1995) emphasize the temporal development of forming impressions about the personality of communication partners in computer-mediated interactions.

Walther and his colleagues argue that people are driven by a strong need for affiliation and for building up impressions of one another. For example, Walter and Tidwell (1995) note that the exchange of information is slower in CMC than in face-to-face interaction, thus the construction of interpersonal knowledge is slower when individuals interact using CMC compared to when they interact face-to-face. According to these researchers, CMC is more task-oriented than person-oriented and less powerful compared to a face-to-face
communication during the early stages of interaction. In this phase, the communication style in CMC situations may be regarded as more informal than in face-to-face interactions, but one is also likely to find cues for tension and emotional evaluations of communication partners or the communicational setting. When CMC participants receive enough opportunities to interact, however, these differences disappear and the communication styles for face-to-face and computer-mediated communication converge. Walther and Tidwell (1995) also emphasize the role of typographical sideways symbols such as emoticons (e.g., :-() or :-(- or :-)- ) or disclaimers (such as *lol*, *yuk yuk*, *rotfl*) as surrogates for nonverbal social cues and they argue that the use of such meta-communicative symbols demonstrates that users adapt the possibilities a medium offers to the requirements of their communication situation.

Thus, Walther (1994) challenges the “filtered-cues” arguments of social presence (Short et al., 1976), social context cues (Sproull & Kiesler, 1986), and media richness theories (Daft & Lengel, 1984; 1986) and characterizes CMC as, in some cases, “hyper-personal” rather than impersonal. For example he cites several studies in which “experienced CMC users rated text-based media, including e-mail and computer conferencing, as ‘as rich or richer’ than telephone conversations and face-to-face conversations” (Walther, 1994, p. 18).

3. Social Presence and Copresence in the Virtual Environments Literature

While communications researchers have focused extensively on identifying media variables that affect the degree to which interactants are salient, or socially present, to one another within technology-mediated interactions, virtual environment and computer science researchers have adopted a somewhat different view of social presence. Whereas the communication studies’ view is that social presence reflects the extent to which a given telecommunications medium can convey the salience of others within a technology-mediated interaction, virtual environment researchers define social presence in a number of different ways.

Heeter (1992) was one of the first in the virtual environment research area to discuss social presence, which she defined as the extent to which other beings, both living and synthetic, exist in a virtual world and appear to react to you. According to Heeter, social presence helps enhance a user’s sense of personal presence within a virtual environment because if other
people are in the virtual world there is more evidence that the world exists and if they respond and react to the user then there is more evidence that the user exists within the world.

Numerous other definitions of social presence have subsequently been proposed within the virtual environment/computer science literature. For example, Biocca and Delaney (1995) defined social presence as the sense of being present in a social encounter with another person, for instance via a telephone call, where two speakers are at different physical locations but can feel fully present with each other in the context of the conversation. Later, however, Biocca (1997) defined social presence as feeling the presence of another intelligent being. McIsaac and Gunawardena (1996) define social presence as the degree to which a person feels "socially present" in a mediated situation while Gunawardena (1995) defines it as the degree to which a person is perceived as a "real person" in mediated communication. IJsselsteijn and his colleagues (IJsselsteijn, de Ridder, Freeman & Avons, 2000) view social presence as the sense of “being together” in a virtual environment, while Lombard (2000) defines social presence as occurring when a person does not acknowledge that technology is creating her/his perception that s/he is communicating with one or more other people or entities.

While these definitions of social presence clearly differ, they seem reflect three primary themes: 1) a user’s sense that there are other social entities co-situated within some interpersonal environment (whether those social entities are other human beings or something else); 2) the extent to which others appear to be real (that is, a user’s sense that they are interacting with a person instead of an agent or some other social entity); and 3) the extent to which social entities are engaged and responsive to one another.

Another way of thinking about this is that the different definitions of social presence reflect three questions a user might ask when they enter an interpersonal environment. First, “Am I alone here, or are others here with me?” Second, “If I’m not alone, who or what is here with me? Is it another person or something else?” And finally, “Whether it’s a person or some other social entity, is it simply here with me or can we interact and engage one another?”

In addition, some views of social presence suggest that users experience a sense of being in the same virtual space with others even when they are, in reality, physically separated. However, this latter view seems confounded with what is sometimes referred to as spatial presence or presence as transportation (Lombard, 2000), which is posited to occur when one
perceives that s/he is in a physical location and environment that is different from her/his actual location and environment in the physical world. Thus, we will not attempt to incorporate this last notion of social presence in our model since it does not appear to reflect anything unique about social interaction per se.

As should now be quite evident, the term social presence is a very fuzzy concept that has been defined in a number of different ways both within and across academic disciplines. This conceptual confusion makes it extremely difficult to integrate the results of studies conducted by different researchers or to draw any kind of meaningful conclusions from the rapidly developing literatures. In addition, the lack of a consensual definition makes it nearly impossible to develop a valid and reliable way of assessing social presence.

Complicating matters further is the fact that other terms, such as copresence and mediated presence are frequently used interchangeably with social presence. For example, Russo (2001) defines mediated presence as the perception by a communicator that another person in a mediated or online environment is “real,” immediate or present. She posits that mediated presence occurs when the communicator perceives the salience and involvement of others and feels that the connection made with another is active, sociable, and sensitive. On the other hand, Slater and his colleagues define copresence as the sense of being and acting with others in a virtual place (Slater, Sadagic, Usoh, & Schroeder, 2000; Durlach & Slater, 2000). Nowak (2001) views copresence as the sense of connection with another mind. And Lombard (2000) argues that copresence occurs when a person does not acknowledge that technology is creating her/his perception that the person or people with whom s/he is engaged in two-way communication is/are in the same physical location and environment when in fact they are in a different physical location. Clearly these definitions of mediated presence and copresence overlap considerably with those put forth for social presence. However, there are no clear lines of delineation between them.

This confounding of terms and lack of consensual definitions makes research in this area an extremely thorny issue. After nearly 30 years of theorizing and research, it appears that no one is clear about what social presence or copresence are, let alone whether or not they contribute to our understanding of technology-mediated social interaction.

An additional issue which has yet to be addressed is whether or not we should be trying to tie social presence or copresence to mediated interactions or, even worse, to a particular
technology, which has been the typical approach taken by researchers in both communications and computer science. Attempting to do so suggests that the perception of others and the psychological process that underlie people’s social responses to them varies across different interactional contexts. There is no reason to believe, however, than humans have evolved different psychological processes to differentiate between these sorts of social contexts. Rather, as evolutionary psychologists have argued (e.g., Kanazawa, 2002), it is more likely that the human brain is hardwired to respond to social stimuli as it did in its ancestral environment, where television, movies, telephones, virtual environments, and other telecommunications media did not exist. Moreover, defining social presence or copresence in terms of either a particular technology or a particular interactional context is extremely limiting in that as the technology or social context changes, the constructs and theory lose all utility. Thus, social presence and copresence may better be viewed as general constructs that operate across interactional settings. However, they are likely to be a number of variables that mediate and moderate the effects of the presence of others within or across these various social contexts.

4. **Differentiating Social Presence and Copresence**

In the next section we introduce our theoretical model of social influence. Before we do so, however, we’d like to propose working definitions of social presence and copresence that are, in part, based on theorizing and research conducted in the sociological and social psychological literatures. Our goal in putting forward these definitions is not to enter into a conceptual debate about how these constructs should be defined. Rather, our intent is to attempt to integrate the different perspectives reviewed thus far, provide definitions that are not tied to a particular technology or interactional context, and to provide some common ground from which to move forward in the remainder of our discussion.

4.1 **Social Presence**

We define social presence as the actual, imagined, or implied presence of others (Allport, 1985) and we argue that others need not be physically present for them to exert social influence over the thoughts, feelings, or behaviors of another. Within the fields of social psychology, communications, and other disciplines, a variety of research methodologies have been used to study peoples’ cognitive, affective, and behavioral responses to others. In some cases, participants interact directly, reflecting actual social presence. In other studies, their
interactions are technologically mediated, reflecting implied social presence. And in still others, participants imagine interacting with others through the use of vignettes or other stimuli.

Fortunately, the results of social influence studies conducted using a variety of methodologies tend to converge, suggesting that social processes operate similarly whether the presence of others is actual, implied, or merely imagined. However, effects sizes for various social phenomena often vary according to the type of social presence manipulation used. For example, face-to-face interactions with confederates or other participants (i.e., actual social presence) tend to be more impactful than are imagined interactions (i.e., imagined social presence).

One possible reason for these differences in effect strength is that the extent to which others are salient in the interpersonal encounter varies depending on whether social presence is actual, imagined, or implied. As the salience of others increases, a perceiver’s awareness of those others is likely to increase as well. Thus, we hypothesize that a sense of copresence (see below) mediates the relationship between social presence (whether actual, imagined, or implied) and social influence outcomes. However, experiencing a strong sense of copresence may not be necessary for all forms of social influence, in that the social presence of others can influence behavior, even when one is not consciously aware that others are present in-person or otherwise. Therefore, we further hypothesize that the relationship between copresence and social influence is moderated by the level of influence response of interest (e.g., low-level, automatic influence processes versus high-level, controlled ones). We will return to these issues later when we introduce our social influence model.

4.2 Copresence

We define copresence as a person’s perception and feeling that others are co-situated within an interpersonal environment. Previously, we suggested that individuals might ask themselves several questions when they enter an interpersonal environment and copresence relates to the first of these, which is, “Am I alone here, or are there others here with me?” Thus, whereas social presence can be thought of as whether or not there are social cues that signify the presence of others within some interactional context, copresence might better be thought of as one’s perception and awareness of those social cues and the corresponding feeling, sense, or belief that others are “there”.

According to this perspective, social presence and copresence are related, yet distinct constructs. Often, social presence leads to a corresponding sense of copresence, which can vary from low (e.g., “I am alone”) to high (e.g., “There are others here with me”). However, this need not be the case since copresence occurs in the mind of the perceiver. For example, there may be, in reality, multiple individuals present within some interpersonal environment (i.e., actual social presence). If an individual, however, lacks awareness of those others, then he may not experience a significant amount of copresence. Similarly, to the extent that various telecommunication media fail to convey the salience of others (i.e., their social presence), interactants’ sense of copresence in the interpersonal encounter will be reduced.

We further suggest that actual social presence is not necessary for interactants to experience copresence. Rather, we believe that copresence can also occur when one perceives that others are co-situated within the same interpersonal environment, even when they are, in reality, physically situated in different locations. In addition, copresence can occur when the others are simply imagined, such as when young children believe that they are in the presence of an imaginary friend. There are, however, likely to be a number of contextual, interpersonal and intrapersonal factors that mediate and moderate the relationship between social presence, copresence, and social influence and we identify and discuss some of these variable below.

In our discussion thus far, we have tried to maintain a conceptual distinction between social presence (which we defined as the actual, imagined, or implied presence of others) and copresence (which we defined as people’s perception of those others and corresponding feeling that others are co-situated within some interpersonal environment) rather than confound them, as has often been the case in previous research and theorizing. In addition, we have purposely maintained a conceptual distinction between copresence and a user’s sense that they are engaged in a social interaction (a feature of social interaction that has frequently been included in previous definitions of social presence or copresence). In the following section we introduce a new construct, social verification, which describes peoples’ perception that they are immersed and engaged in an interaction with another. We have chosen to distinguish between copresence and social verification, however, because we feel it is important to differentiate between social influences that result from mere copresence versus those that result from social interaction.
5. A Model of Social Influence

In this section, we introduce our model of social influence. We begin by proposing a general model that attempts to describe the social dynamics of social influence regardless of the interaction context. Next, we build on this model to discuss factors relevant to a variety of technology-mediated interactions. Finally, we discuss factors that are unique to social interaction and social influence in collaborative virtual environments (CVE).

5.1 Social Influence

Social psychology has been defined as the scientific discipline that attempts to understand and explain how the thoughts, feelings, and behaviors of individuals are influenced by the actual, imagined, or implied presence of others (Allport, 1985). As this definition suggests, individuals are both directly and indirectly influenced by the presence of other people and social influences range from those that are subtle and covert to direct and overt attempts to influence feelings, attitudes, beliefs, and behavior.

The social psychological research literature is replete with reports of how the presence of others triggers a host of affective (e.g., feelings, preferences, prejudices), cognitive (e.g., attributions, impressions, attitudes, beliefs, stereotypes, social categorization), and behavioral (e.g., performance facilitation or inhibition, discrimination, conformity, compliance, obedience, helping, aggression) social responses. In some cases, the mere presence of others is sufficient to activate such reactions in the absence of any meaningful social interaction, even when that social presence is simply imagined or implied. In other instances, however, full-blown social interaction is necessary. Thus social presence and/or the resulting sense of copresence alone are insufficient explanations for the full range of social influence phenomena that have been documented. Thus we believe that it is important to not only examine the extent to which one perceives that he or she is in the copresence of others, but that it is also important to assess the extent to which one feels embedded and engaged in a meaningful or symbolic social interaction (Mead, 1934).

5.2 Social Verification

A key component of our model of social influence is social verification, which reflects the extent to which individuals experience interpersonal encounters in ways that verify that they are engaged in meaningful interpersonal communication (i.e., symbolic interaction, Mead,
Perceiving and Responding

According to our conceptualization, one’s subjective sense of social verification can vary continuously from low to high. At the low end of the social verification continuum, one may be aware of the presence of others (and may, thus, experience a sense of copresence), but he does not feel engaged in any kind of interaction with them. At the high end of the social verification continuum, on the other hand, one not only experiences a strong sense of copresence, but she also feels deeply involved and embedded in a meaningful interpersonal interaction. Thus, social verification relates to the third question we suggested people might ask when they enter an interpersonal environment, namely, “Whether it is a person or some other social entity, is it simply here with me or can we interact and engage one another?” (We will address the issue of whether an individual believes that she is interacting with another person or with some other social entity later when we discuss social interaction with agents vs. avatars in CVEs).

As an example, imagine a classroom full of students taking an exam. In this situation, the students’ attention is probably focused extensively on the task at hand (i.e., completing the exam). While taking the test, however, students may simultaneously maintain awareness that they are not alone in the environment and this awareness may have consequences for their behavior (e.g., a student’s awareness of the professor may discourage him from cheating). Despite this awareness of others, however, students are not likely to perceive that they are engaged in an interpersonal interaction. In fact, interacting with others in this situation, even at a minimal level, would be strongly discouraged! Thus, in this example, a given student may simultaneously experience a moderate to high sense of copresence (i.e., he is aware that he is not alone), but a low sense of social verification (i.e., although he is not alone, he is not interacting with anyone).

Similarly, low levels of social verification may result when one is performing some solitary activity in a public location, such as when someone rides her bike along a public trail. When one is riding, her attention may be focused on avoiding obstacles, selecting the route, enjoying the scenery, listening to music on a portable stereo, or her mind may be preoccupied with daydreams or other thoughts. Although the rider may also maintain some awareness of other people in the periphery (reflecting a low to moderate sense of copresence), there is likely to be little or no interaction between individuals. Nonetheless, the presence of others may influence the rider’s behavior. For example, research on social facilitation (e.g., Triplet, 1898; Zajonc, 1965) demonstrates that the mere presence of others may facilitate
performance on an easy or well-practiced task. Thus, the rider may find herself riding faster when others are present, even though her sense of both copresence and social verification are relatively low.

At a somewhat higher level of social verification, one may not only be aware of the presence of others, but he may also perceive that this awareness is reciprocal. Various nonverbal and verbal behaviors such as establishing mutual eye gaze, smiling, nodding, saying “hello,” etc., may convey a sense of mutual awareness. While there is clearly some interpersonal interaction involved here, the level of interaction is minimal. Moreover, while the exchange of these nonverbal and verbal behaviors may be sufficient to induce a relatively strong sense of copresence, social verification would likely be only of moderate intensity.

At the high end of the social verification continuum, on the other hand, an individual would feel immersed and engaged in a meaningful social interaction, for example when two individuals are deeply engaged in conversation with one another. Such interaction is likely to involve both verbal and non-verbal elements and is apt to be characterized by feelings of interpersonal connectedness, intimacy, emotional involvement, a sense of shared meaning, and the perception of shared reality. In such a case we would expect both copresence and social verification to be high.

5.2.1 Copresence and Social Verification

Although we believe that copresence and social verification are positively related, we feel that it is important to maintain a conceptual distinction between them and to assess them separately in empirical investigations. While we would expect one’s subjective sense of copresence to increase as social verification increases, we also believe that it is possible to experience a high level of copresence while social verification is low. In addition, we hypothesize that copresence and social verification will differentially predict various social influence outcomes. In some cases, copresence alone may be sufficient for social influence. However, in other instances, we would expect that both a high level of social verification and a strong sense of copresence are necessary for social influence to result.

5.3 Social Influence Response System

As noted previously, the mere presence of others (reflecting low social verification and perhaps even low copresence) is sometimes sufficient to induce social influence. Other
times, however, individuals must experience a high degree of social verification for social influence to occur (i.e., they must perceive that they are involved in a meaningful social interaction with an attentive and responsive communication partner). We hypothesize that the level of social verification necessary to produce social influence is moderated by the response system used to assess the operation of social influence effects.

For low-level response systems, such as unconscious reflexes, actions, and other automatic processes, we hypothesize that social presence may lead to social influence, even when social verification and perhaps even copresence are low. In fact, low-level social influence may occur even when one is only minimally aware of the presence of others. For high-level response systems, such as the ones involved in conscious, controlled, and purposeful actions or verbal communication, however, we would expect there to be a strong, positive relationship between social verification and social influence.

5.4 Factors Affecting Social Verification & Copresence

The extent to which the actual, imagined, or implied presence of others (i.e., social presence) leads to a sense of copresence and/or social verification is likely to be a function of a number of contextual, interpersonal, and intrapersonal factors.

5.4.1 Contextual Factors

Contextual factors include elements of the social context as well as other features of the interactional situation that affect the extent to which individuals are salient to one another. Noise and other environmental distractions direct individuals’ attention away from each other, thereby reducing their sense of copresence with one another and their ability and motivation to interact. Similarly, copresence and social verification may vary depending on the type of social presence (i.e., actual vs. imagined vs. implied). However, we hypothesize that there are a number of interpersonal and intrapersonal factors that moderate this effect so that comparable levels of copresence and social verification are possible for actual, imagined, and implied social presence.

5.4.2 Interpersonal Factors

Interpersonal factors include the relationship among the interactants; their relative status to one another; their interaction history; their attitudes, beliefs, and expectations about one
another; and the extent to which they perceive each other to be attentive, involved, and responsive within the interaction. Both verbal (e.g., appropriately timed and synchronized responses) and nonverbal behaviors (e.g., eye contact, head nods, smiling and other facial expressions, body orientation and position, and vocal expressiveness) are essential to conveying liking, attention, involvement, and responsiveness and are, thus, vitally important in producing a sense of social verification.

5.4.3 Intrapersonal Factors

Finally, intrapersonal factors, including motivational (e.g., self-relevance, interaction goals), affective (e.g., mood), dispositional (e.g., personality, temperament, intelligence), psychological (e.g., opinions, beliefs, attitudes, cognitive resources), organismic (e.g., age, sex, health), and demographic characteristics (e.g., race, socio-economic status, religion), may also affect the extent to which individuals experience a sense of copresence and/or social verification.

The relationship between these contextual, interpersonal, and intrapersonal factors is hypothesized to be highly complex and it is expected that they may often interact when contributing to both copresence and social verification. In addition, although we’ve identified a number of factors that may affect the extent to which copresence and/or social verification are experienced in an interpersonal encounter, we by no means wish to imply that this list is exhaustive. On the contrary, we would encourage others to identify additional factors that may affect how individuals perceive and respond to one another and to conduct empirical research designed to understand how these sorts of factors contribute to one’s sense of copresence and social verification in a variety of interactional contexts.

6. Social Influence in Technology-Mediated Interactions

Up to this point, we have discussed general factors that we believe to be of import to understanding social interaction and social influence across a variety of interactional contexts. However, there are some additional factors that may affect the extent to which one experiences copresence and social verification in technology-mediated interactions and which may have implications for the extent to which various technologies can be used successfully for social interaction and social influence.
Although we do not believe that mediated interactions are inherently different or inferior to non-mediated ones, there are a number of dimensions along which communications media vary, including whether communication occurs synchronously or asynchronously, how many channels of social information they convey, the number and types of social context cues transmitted, the mode(s) of communication (e.g., text, auditory, visual), their level of interactivity, and their degree of immersiveness. These and other factors may affect the way in which individuals’ perceive one another and their ability to interact. However, we would expect a number of other interpersonal and intrapersonal factors, such as a user’s experience communicating via a particular medium, the interactants’ relationship to one another, or their interaction goals, to moderate these effects.

All else being equal, we would expect media that do a better job of conveying the salience of others (i.e., implied social presence) to produce a stronger sense of copresence than will media in which others are less salient. Thus, consistent with the communication studies view of social presence, we would expect one’s sense of copresence to be stronger when interacting over the telephone (which conveys a broad range of verbal and non-verbal auditory information) than one would experience when interacting via a text-based instant messaging system.

In addition, we expect that media that convey more authentic and reliable social information and that allow individuals to interact naturally and in real time, will produce a greater sense of social verification than will media that are less authentic, unnatural, or that convey less social information. However, we would suggest that it is not only the amount of social information transmitted that matters. Rather, the information that is communicated must also be veridical and non-distorted. When social signals become distorted, or qualities of the media (such as lag) disrupt normal information exchange, then social verification is likely to decrease and this will affect the outcome of the interaction.

Anyone who has used a videoconferencing system (especially older ones) knows that interactions are often uncomfortable because non-verbal signals, such as eye gaze, become distorted when the number of interacts exceeds two (and even in a two-way interaction it is difficult to calibrate the cameras so that mutual gaze is properly rendered). In addition, system lag can produce discrepancies between auditory and visual signals which can hinder interaction. This is arguably why many prefer to use a telephone rather than a
videoconferencing system, even though the telephone conveys less social information (i.e., only auditory signals rather than both visual and auditory signals).

7. **Social Influence in Collaborative Virtual Environments**

The use of virtual environments for collaboration and social interaction offers a number of potential advantages over other telecommunications media. As with most other telecommunication technologies, users can interact with others who may not be physically present in the "real" environment. In addition, both physically present and remote users can be co-situated within the same virtual environment though the use of an augmented reality system.

Because of its immersive, interactive nature and its ability to render both auditory and visual signals (and perhaps haptic and olfactory signals in the future), a greater variety of verbal, non-verbal, and contextual social information can be exchanged using CVEs than is possible with many other telecommunications systems. In addition, social context cues can be provided by modeling certain aspects of the physical environment, by transmitting information about the relative status of the actors, and by conveying non-verbal behaviors that define the nature of the social situation (e.g., turn-taking behaviors that regulate the social interaction). Additionally, a variety of environmental distractions and noise can be removed (i.e., visual and auditory information that does not relate to the social interaction or social context can be filtered out) through the use of an immersive virtual environment system (e.g., either a CAVE-like room in which a user is isolated and immersed in a virtual environment or through the use of a head-mounted display).

7.1 **Embodiment in Virtual Environments**

In contrast to most other forms of technology-mediated interaction, social interactants are represented or “embodied” within a virtual environment by an avatar (i.e., a graphical representation of the user). Embodiment is important because it enables CVE users to convey important social information about themselves. In many CVE applications, users can create their own unique avatar or they can choose and customize an avatar from an extensive library. Sometimes, users create or select avatars that resemble them or match them on a variety of socially-relevant dimensions such as age, status, race, gender, physical appearance, etc. However, individuals can also use their avatar to create a new social identity.
The avatar that an individual creates or chooses can convey a great deal of social information and avatars can vary along a number of dimensions in terms of their appearance and behavioral capabilities. These characteristics may affect how individuals perceive and respond to one another within CVEs.

7.1.1 Static versus Dynamic Avatars

Avatars can be static or dynamic. Static avatars are placed in a virtual environment to convey the presence of others, but the avatars do not move or interact in any way. Dynamic avatars, on the other hand, may be moved around the virtual environment to convey information about the proximity and orientation of users. In addition, dynamic avatars may possess the capability to display a variety of behaviors (such as gestures, expressions, etc.) that can be used to communicate (see behavioral and social realism below).

7.1.2 Dimensionality

Avatars can be embodied either be two-dimensional photographs or drawings or three-dimensional characters or objects. Two-dimensional avatars may be either fully static (see above) or they might be moved around the virtual environment to indicate the user’s current location and orientation. Three-dimensional avatars, on the other hand, may possess more dynamic capabilities.

7.1.3 Adornment/Customizability

The avatars that individuals use within a given application can either come from a standard avatar library or they may be customized or created by the user. Avatars that are customizable or user-created may convey more social information than do generic avatars. For instance, the way in which a user chooses to dress or adorn his or her avatar may be socially meaningful and relevant. However, even when a generic avatar library is used, the particular avatar selected may be socially meaningful provided the user was able to choose his own avatar rather than have it assigned to him.

7.1.4 Anthropomorphism

Anthropomorphism refers to the extent to which an avatar resembles a human being (Nowak, 2001; Nowak & Biocca, 2001). Theoretically, a user can choose anything to represent him- or herself within a CVE. Thus avatars can range from inanimate objects (including simple
geometric shapes), to animals, to imaginary characters (e.g., aliens, fairies, and other mystical creatures) to human-like figures. Avatars that are highly anthropomorphic look like human beings while avatars that are low in anthropomorphism may include animals or inanimate objects. Because people have a tendency to endow animals with human-like qualities, however, animal avatars may be viewed as more anthropomorphic than are avatars comprised of inanimate objects.

7.1.5 Photo-Realism

Photo-realism refers to the extent to which an avatar appears cartoonish versus photo-realistic. Recent advances in avatar-creation software (e.g., 3D Me Now software) have made it quite easy to create photographically realistic avatar heads that closely resemble the human user. Modeling realistic-looking human bodies is more difficult, however, primarily because low polygon count models are needed to keep frame rates at an acceptable level. Thus, avatars might have a photographically realistic head paired with a less-realistic looking body.

Many people tend to think of virtual environment technology as a visual medium, and, thus, might be tempted to assume that it is important to represent objects and people as photo-realistically as possible. Bailenson and his colleagues (Bailenson et al., 2001) have found, however, that photographic realism makes very little difference in terms of how people perceive and respond to 3D characters, at least those that are anthropomorphically embodied.

7.1.6 Behavioral Realism

More important than photo-realism (Bailenson et al., 2001), and perhaps even anthropomorphism, is an avatar’s behavioral realism. Behavioral realism refers to the extent to which avatars and other objects in a virtual environment behave like their counterparts in the physical world (Blascovich et al., 2002). Cartoonists have long known the importance of behavioral and interactional realism (see below) and have used it to devise compelling characters whose cartoonish, non-anthropomorphic appearance (e.g., mice, ducks, pigs) nonetheless creates the illusion of life.

Behavioral realism is usually accomplished through the use of motion capture hardware or special computer animation software. In general, behavioral realism is achieved when an avatar’s behavior resembles that of its non-virtual counterpart. For example, a human-like
avatar would be behaviorally realistic when its movements appear to be generated by an actual human being.

However, an avatar may also be considered behaviorally realistic when its movements and behaviors are “believable,” even when they do not match those of the avatar’s real-world counterpart. Thus, for example, a cat-like avatar may be behaviorally realistic either when it moves or behaves like a typical cat or when it moves or behaves in a way that creates the illusion of sentience, personality, and intelligence. For example, the Disney character Mickey Mouse did not act like a typical mouse. Rather, his movements and behaviors were more like those of a human being. Nonetheless, he would have been considered to be a behaviorally realistic character because his movements and behaviors were so believable.

7.1.7 Interactional Realism

Finally, avatars may vary in terms of their interactional realism. Interactional realism refers to the extent to which avatars are capable of responding in socially meaningful and appropriate ways. Interactional realism can range from low to high, with low interactional realism being reflected by a static, non-interactive avatar and high interactional realism being reflected by a dynamic, interactive, and potentially responsive one.

7.1.8 Avatar Characteristics, Copresence, Social Verification, and Social Influence

In the context of technology-mediated social interaction, we predict that, all else being equal, copresence, social verification, and social influence will be positively related to an avatar’s characteristics. In other words, we expect copresence, social verification, and social influence to be greatest when dynamic, three-dimensional, customizable, anthropomorphically embodied, and photographically, behaviorally, and interactionally realistic avatars are used.

Consistent with this idea, Nowak and her colleagues (Nowak, 2001; Nowak & Biocca 2001) demonstrated that level of anthropomorphism affects people’s social judgments of their interaction partners. In one study, interaction partners represented by more anthropomorphic images were rated as more socially attractive and they received higher partner satisfaction ratings than did partners represented by either no image or by a low-anthropomorphic image (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).

Although of all of these dimensions along which avatars vary may affect how individuals perceive and respond to avatars with CVEs, we believe that interactional realism is the most important and that the interactional realism of an avatar may moderate the effect of the other dimensions on copresence, social verification, and social influence. In particular, we predict that anthropomorphism and photo-realism will interact with the interactional realism of the avatar such that differences in copresence, social verification and social influence will be greatest between high and low anthropomorphic and/or high and low photographically realistic avatars when interactional realism is low. At high levels of interactional realism, however, we expect many of these differences to disappear. Thus, we expect that non-anthropomorphic, cartoonish avatars may produce both high levels of copresence and social verification, leading to social influence comparable to that of highly anthropomorphic and photographically realistic avatars, provided they exhibit sufficient interactional realism.

Moreover, we predict that differences in copresence, social verification, and social influence between actual (e.g., face-to-face) and implied (e.g., technology-mediated) social presence will be moderated by an avatar’s interactional realism. Thus, we expect interactional realism to affect the extent to which others are perceived as salient, responsive, and involved in an interaction, and this will then enhance copresence, social verification, and social influence. However, consistent with our model of social influence (see also Blascovich et al., 2002, Blascovich, 2002a; 2002b) we expect that these effects will be further moderated by the social influence response level examined. Thus, in our view, greater interactional realism is required to produce high-level compared to low-level social influence.

8. Human-Computer Social Interaction

Up to this point, we have limited our discussion to human-human social interaction. First, we proposed a general model of social influence that describes a variety of factors that we believe affect the extent to which the presence of others leads to various forms of social influence across a variety of interactional contexts. Next we discussed addition factors that may influence the way in which individuals perceive and respond to others in technology-mediated social interactions. Finally, we discussed human-human social interaction within collaborative virtual environments and we reviewed some of the ways in which avatars vary
and how these avatar characteristics may affect people’s perceptions of one another and their social responses.

While a great deal is now known about the social processes involved in human-human social interaction and social influence, much less is known about the social dynamics of human-computer social interactions. Some researchers have suggested that individuals unconsciously attribute human characteristics such as gender or ethnicity to computers and other interactive media representations and that they apply social rules and expectations when interacting with them (e.g., Lee & Nass, 2001; Moon & Nass, 1996; Nass & Moon, 2000; Nass et al., 1995, Nass et al., 1994; Nowak, 2001; Nowak & Biocca, 2001; Reeves & Nass, 1996). However, not all research has supported this view (e.g., Bailenson et al., 2001; Bailenson et al., in press; Blascovich, 2002b; Blascovich, et al., 2002; Hoyt et al., in press; Swinth & Blascovich, 2001). Therefore, it is important to understand how human-computer interactions are similar to as well as different from various types of human-human interactions and to identify factors that may moderate peoples’ social responses to computers and other synthetic social entities.

8.1 Interactive Software Agents

In the not so distant past, computers were viewed primarily as productivity tools and a lot of human-computer interaction research focused on improving the interface so that computers would be easier to use and so that users could be more productive in their interactions with them. More recently, interactive software agents, such as the infamous Microsoft® paperclip were introduced to try and make the computing experience more sociable and enjoyable. While these early agents failed to gain the initial popularity that software designers might have hoped, the use of computerized agents has continued to grow in recent years with agents serving as interactive guides on a number of websites and in a variety of educational, entertainment, and productivity software applications.

Unlike the first generation agents which were often nothing more than a new face on a simple search engine, many of today’s agents are capable of autonomous decision-making and even learning or adapting to their environment (in which case they are called “autonomous agents” [Franklin, 1997] or “intelligent agents” [Hedberg, 1996]). While not all agents are “intelligent” (many are fully scripted to carry out specific tasks or behaviors) most of the agents used in internet and software applications today are endowed with distinct personalities and many are capable of communicating using a natural language interface.
Similarly, the use of agents as social entities within virtual environments has also grown in popularity in recent years and these agents can be programmed to display humanlike intelligence and behavior, especially in contextually-constrained interactions. Because of recent advances in artificial intelligence and computer graphics technology accompanied by a rise in the use of avatars to represent human users in CVEs (Damer, 1997), users find it increasingly difficult to tell whether the social entities they encounter in a CVE (or some other technology-mediated interaction) represent other real people or not. Thus, it is often difficult for them to determine whether they are engaged in a human-human or a human-computer social interaction (Brent & Thompson, 1999; Dryer, 1999).

### 8.2 Embodied Agents

As discussed previously, we use the term “avatar” to describe a graphical character or object that represents a human being within a collaborative virtual environment. In contrast, we call a graphical character or object an “embodied agent” if it is controlled by or represents an agent (i.e., an artificial intelligence algorithm) rather than another person. Thus an avatar is the embodiment or representation of a human interactant while an embodied agent is a graphical representation of a computer program that has been designed to interact with, or on behalf of, a human being. While the primary difference between avatars and embodied agents is in whether the entity or “intelligence” behind the representation is a human (i.e., avatar) or a computer program (i.e., embodied agent), in reality control over a virtual representation can switch back and forth from human to computer or some behaviors can be controlled by a human while others are controlled by a computer.

Thus, in our view, how a virtual representation is perceived by a human user is more important than the actual agency of the virtual embodiment (i.e., whether the volitional or intentional force that actually drives the actions of the embodiment is a person or a computer or some combination). To the extent that people perceive that they are in the presence of and interacting with an intelligent, sentient, interactive, and socially responsive entity (regardless of whether they are, in fact, engaged in a human-human or human-computer interaction), we expect their behavior and responses to resemble those elicited during normal human-human social interaction. However, to the extent an embodied agent fails to provide convincing evidence that it is aware of a human user’s presence and fails to interact and respond appropriately, we expect the embodied agent’s ability to exert social influence to be reduced.
8.3 Embodiment Characteristics of Agents

One of the primary purposes of embodiment is to convey the (implied) social presence of others within a virtual environment. As we discussed previously, graphical characters can vary along a number of dimensions and because agents and people can be embodied or represented in the same way within a virtual environment, we expect these embodiment characteristics (with the exception of customizability) to affect how individuals perceive and respond to embodied agents, just as they are hypothesized to influence peoples’ perceptions and responses to avatars. In other words, the more realistic an agent’s appearance and behavior, the more we expect people to perceive and respond to an embodied agent as they would another real person.

8.4 Social Interactions with Embodied Agents versus Avatars

In contrast to what others have suggested (e.g., Nowak, 2001; Nowak & Biocca, 2001, Reeves & Nass, 1996), we believe that peoples’ knowledge that social entities within a virtual environment represent agents rather than other people can have consequences for how they perceive and respond to the implied presence of others. Specifically, according to our model of social influence within virtual environments (Blascovich, 2002a; 2002b; Blascovich, et al., 2002) we expect that, all else being equal, users should experience lower levels of both copresence and social verification when they are cognizant of the fact that they are in the presence of and/or interacting with an embodied agent compared to an avatar. Consequently, we hypothesize that the general ability of embodied agents to exert social influence over the thoughts, feelings, or behaviors of human users should be less than that of avatars.

In reality, however, all things are rarely equal. Thus, we believe it is theoretically possible for embodied agents and avatars to elicit similar levels of copresence, social verification, and social influence. Although we have discussed elsewhere the situations when we believe human-human and human-computer interactions are most likely to differ (see Blascovich et al., 2002; Blascovich, 2002a; 2002b), one of the factors that is important to consider is the type of response one wishes to assess. With low-level, automatic, responses (such as maintaining appropriate interpersonal distance or conforming to group norms), one’s beliefs about whether one is interacting with an agent or an avatar appears not to matter very much (Bailenson et al., 2001; Bailenson, et al., in press; Swinth & Blascovich, 2001). However,
when dealing with higher-level, controlled processes, one’s beliefs about whether one is interacting with an agent or an avatar may matter a great deal (Hoyt et al., in press).

In addition, responses to embodied agents and avatars may differ depending upon the interpersonal self-relevance of the interaction. According to our model of social influence (Blascovich, 2002a; 2002b; Blascovich et al., 2002), the greatest differences in perceptions and responses to embodied agents and avatars should be expected when an interaction involves core beliefs or attitudes or when strong emotions are involved (i.e., the interaction holds a great deal of interpersonal self-relevance). When interpersonal self-relevance is high, a strong sense of both copresence and social verification may be required for social influence to occur. Because we expect people to typically experience greater copresence and social verification when they feel they are interacting with avatars compared to embodied agents, it should be easier for an avatar to induce social influence than an embodied agent and, in fact, an embodied agent may need to be extremely realistic looking and behaving to elicit social influence. In contrast, whether one is interacting with an embodied agent or avatar is unlikely to make much of a difference for interactions that are low in interpersonal self-relevance.

Finally, as we have discussed elsewhere (Blascovich, 2002a; 2002b; Blascovich et al., 2002), when users believe they are interacting with embodied agents rather than avatars, some level of realism (most importantly, interactional realism) is necessary for social influence and meaningful social interaction to occur (see for example, Pertaub, Slater, & Barker, 2002). Moreover, the amount of realism required is a function of the level of behavior one wishes to assess, with greater realism required for higher-level behaviors. In interactions involving avatars, however, realism appears not to matter as much, since the mere belief that one is interacting with another person may be sufficient to induce thoughts, feelings, and behaviors that are characteristic of typical human-human interactions (e.g., see Bailenson et al, 2001; Swinth & Blascovich, 2001).

9. Conclusion

Research on technology-mediated social interaction is currently a hot topic. Following the tragic events in the United States on September 11, 2001, the growing trend of using telecommunications technology for collaborative meetings and social interaction is likely to continue, bringing with it an urgent need to understand more thoroughly how users of various
Perceiving and Responding

telecommunication technologies experience their mediated interactions with others. However, conceptual ambiguities and the confounding of key concepts such as social presence and copresence have made it difficult for researchers to integrate the results of the rapidly expanding research literature on technology-mediated interaction or to draw any meaningful conclusions from the assortment of empirical studies that have been conducted. It is, therefore, our hope that this paper will provide some clarification and theoretical direction to help guide future investigation of social interaction in both technology-mediated and non-mediated interactional contexts.
10. References


