Chapter 5

Transformed social interaction in mediated interpersonal communication

Jeremy N. Bailenson, Nick Yee, Jim Blascovich, and Rosanna E. Guadagno

Background

Over time, our mode of remote communication has evolved from written letters to telephones, email, internet chat rooms, and video-conferences. Similarly, virtual environments that utilize digital representations of humans promise to further change the nature of remote interaction. Virtual environments are systems which track verbal and nonverbal signals of multiple interactants and render those signals onto *avatars*, three-dimensional, digital representations of people in a shared digital space. Unlike telephone conversations and video-conferences, interactants in virtual environments have the ability to systematically filter the physical appearance and behavioral actions of their avatars in the eyes of their conversational partners, amplifying or suppressing features and nonverbal signals in real time for strategic purposes. These transformations can have a drastic impact on interactants' persuasive and instructional abilities. Furthermore, researchers can use this mismatch between actions performed by a speaker and actions perceived by an audience as a tool to examine complex patterns of nonverbal behavior which are difficult to isolate in face-to-face interaction.

We first discuss a framework for classifying digital human representations and the role they play in Computer Mediated Communication (CMC). We then present a theory called Transformed Social Interaction (TSI) that explores how CMC allows people to interact in ways not possible face-to-face. We review a number of published studies examining TSI as well as summaries of new, unpublished data and work that is currently in progress. We conclude by relating CMC to theories of social influence, discussing the next step in digital human research and applications, and discussing potential ethical problems with TSI.

A framework for digital human representation

The study of digital human representation within CMC has progressed significantly over the past 15 years, including conceptual, design, and empirical issues. Currently, vast numbers of individuals interact with digital versions of each other on at least a daily basis. The digital human forms utilized during these interactions range from digital audio representations on cellular phones to icons within emails to graphical representations in video games and chatrooms. In this section, we describe research approaches that provide frameworks relating to virtual humans.

Traditionally, researchers have distinguished *embodied agents*, which are models driven by computer algorithms, from *avatars*, which are models driven by humans in real time. Most behavioral research examining social interaction between people and virtual humans has utilized embodied agents (as opposed to avatars—see Bailenson & Blascovich, 2004, for a discussion). One reason for this disparity is that readily available commercial technology allowing individuals to create digital avatars which can look like and behave in real time like the individual has emerged only recently. Previously, producing real-time avatars that captured the user's voice, visual features, and subtle movements was quite difficult. Consequently, understanding the implications of the visual and behavioral veridicality of an avatar on the quality of interaction is an important question that has received very little empirical attention (see Schroeder, 2002, for a review of the existing empirical work on avatars).

Avatars are digital models that may look or behave like the humans they represent. In virtual environments, avatars are often rendered dynamically, in real time, to reflect at least some user behavior or movements (e.g., Reidsma *et al.*, 2005). However, when applied to more traditional forms of CMC, the definition of an avatar is fuzzy. For example, the definition of avatar including "looking like a user" encompasses a digital photograph, such as one posted on an online dating website. Some would object because such an image has little or no potential for behavior or movements. However, others would argue that people utilize static (i.e., nonanimated) avatars in synchronous internet chat. While many discuss the concept of avatars in the CMC literature, a standard definition of avatars that researchers subscribe to has not emerged. Here, we believe it important to examine the suitability of different types of avatars for representing the user (Konijn & Hoorn, 2004).

Figure 5.1 provides a preliminary framework for considering representations of humans. The abscissa for each graph represents form similarity, how much the representation statically resembles features of a given person. The ordinate for each graph denotes behavioral similarity—how much the behaviors of the representation correspond to the behaviors of a given person. The graph on the left classifies representations that correspond to a given person's form or behavior synchronously or in real time. The graph on the right classifies representations that correspond to a person's form or behavior asynchronously.

Illustrating synchronous avatar behavior (left side of Figure 5.1), a

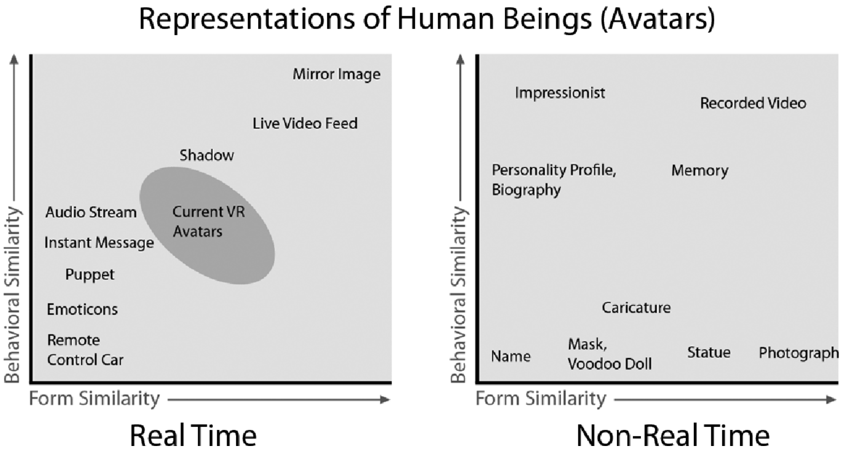

Figure 5.1 A framework for classifying representations of humans in physical and digital space

puppet is a representation of a person that has high behavioral similarity (the movements of the puppet's mouth are closely tied to the person controlling it) but low form similarity (a puppet does not look like the person controlling it). However, the controlling person's behaviors are expressed in real time. On the other hand, an impressionist (i.e., someone who can very closely reproduce or mimic the behaviors of a person who is not physically present) has high behavioral similarity by definition, but only high form similarity if the impressionist actually looks like the person being mimicked. Unlike the puppet, however, the impressionist is typically an asynchronous representation—the person being mimicked need not be present, aware of the impressionist's existence, or even still alive.

As Figure 5.1 demonstrates, representations of human beings can take many forms. The shaded oval denotes the space in which we typically discuss avatars—digital representations of humans that are utilized in virtual environments. Blascovich *et al.* (2002) provide a theoretical framework to determine the interplay of behavioral and form realism for the avatars that fall into this shaded region.

Digital humans today

Currently, digital humans are used in a number of CMC venues. For example, sound is transformed into digital information as it travels over fiber-optic cables and cellular networks; consequently, the audio representation we perceive over phone lines is actually an acoustic avatar of the speaker. This classification may seem trivial at first, but becomes less trivial when pres*et al*gorithms are applied to the audio stream to cause subtle changes in the acoustic avatar (e.g., Nass & Brave, 2005), such as cleaning and amplifying the signal or making phonetic assumptions concerning specific languages. In other words, because the voice is translated into digital information, it is an abstracted representation of the human, as opposed to raw perceptual input from the speaker.

The internet is filled with different forms of CMC employing digital representations. For example, as two people communicate via instant messaging (IM), they appear to each other as a series of text messages, emoticons, and pauses. Recent estimates show that 53 million people in the United States use IM (Project, 2004). Of those born after 1976, 62 percent use IM on a regular basis. Furthermore, there is ample use of video-conferencing technology; with digital video one can consider a human representation an avatar. Moreover, digital representations are also seen in video games. Currently, about 50 percent of the United States' population plays video games (ESA, 2005), men and women alike. On average, gamers spend about 7.5 hours per week playing (ESA, 2005).

Perhaps the best example of social interaction via graphical digital representation occurs in a genre of video games known as massively multiplayer online games (MMOGs). Millions of players spend on average 22 hours a week interacting, collaborating, and competing with each other via graphical avatars (Woodcock, 2005; Yee, in press). Users are often given a great degree of control over the appearance of their avatars. For example, in the game Star Wars Galaxies, users can alter their avatar's gender, age, height, weight, musculature, hair style and color, eye shape and color, lip fullness, cheek fullness, nose protrusion, freckles, baldness, and so on.

Human desire to transform representation

According to Goffman's approach to understanding identity, the presentation of the self must be understood as a constant performance in front of the social audience around us—that we choose our gestures, mannerisms, and actions to give off a desired impression of the self to others (Goffman, 1959). Of course, psychologists have also long noted that the primary function of self-presentation is instrumental. These forms of presentation involve making favorable impressions on others to gain social advantage, such as appearing pleasant or likeable (Jones, 1964; Jones & Pitman, 1982). In general, people want to present themselves positively and strike a balance between favorability and plausibility (Schlenker, 1980).

Alterations to self-presentation occur in many different forms and many different ways (see Figure 5.2). A wide range of cosmetic products and services provide short-term enhancements to our appearances. These include makeup, haircuts, and hair styling products among others. We also alter our nonverbal behaviors consciously and unconsciously for social advantage. For example, mimicking another person's gestures and behaviors

	Appearance	Nonverbal Behavior	Verbal Behavior
Short Term	Haircuts Makeup	Mimicking Ingratiating Gestures	Lying Word Choice
Long Term	Plastic Surgery Dieting	Habit Suppresion Table Manners	Oratory Training Language Acquisition

Figure 5.2 Methods utilized to transform physical representations

for social rapport can occur both automatically and in a controlled fashion (Chartrand & Bargh, 1999).

It also occurs via verbal behavior, such as illustrated by speech accommodation theory—the process by which our accent, speech rate, and lexical choices come to converge with people we talk to (Giles & Claire, 1979).

Alterations to self-presentation can also be long term or even permanent. For example, plastic surgery or weight training can provide more long-term effects on our appearances. Also, we learn a variety of nonverbal social rituals as children, such as table manners, and learning a new language can be viewed as a form of long-term alteration to our verbal behavior.

Transformed social interaction

In CMC, it is relatively trivial for a person to transform many aspects of their own avatar as well as the social world in which they interact. Consider the depiction of CMC depicted below in Figure 5.3.

The right panel indicates that the behaviors of three users in separate remote physical locations are tracked. Various technologies can be used to track various features of the users, such as voice, appearance, and movements (e.g., facial expressions, gestures). The left panel demonstrates a digital configuration in which the three users are rendered in the same virtual location where they can see and hear each other's avatar. Such virtual locations can be a teleconference, video-conference, chatroom, video game, or an immersive collaborative virtual environment.

In many CMC systems, each user has a digital image of the others' avatars stored locally on his or her system. The system receives digital

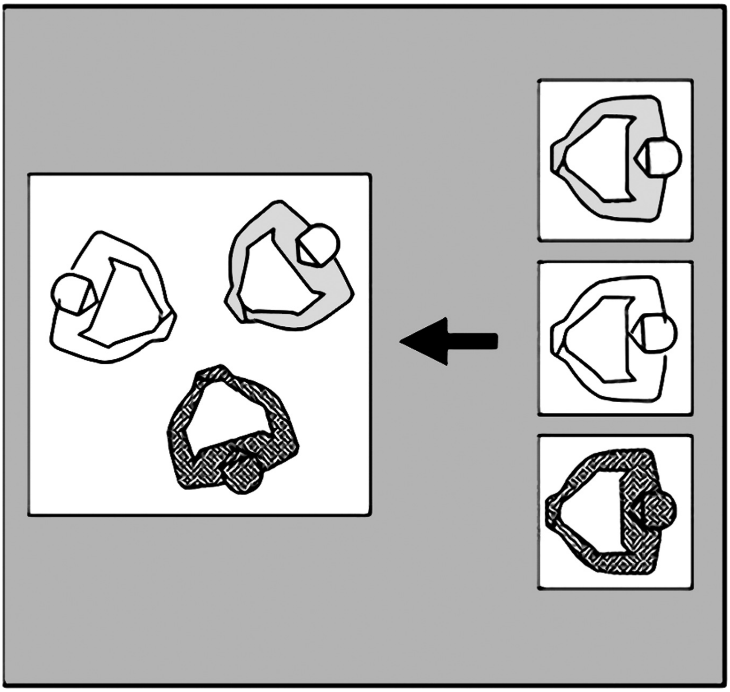

Figure 5.3 A schematic of a simple CMC using digital human avatars

tracking information regarding voice, movements, gestures, and other actions over a network, allowing the dynamics of the avatars stored in his or her system to be updated continuously and rendered more or less veridically. Given that CMC systems must render the world separately for each user simultaneously, it is possible to render the avatars differently for each user at the same time. In other words, for each CMC user, tracking devices transmit a stream of information that indicates his or her actions. However, that stream of information can be altered in real time for strategic purpose by system operators, who may or may not be the users themselves. The theory of Transformed Social interaction (TSI, see Bailenson, 2006; Bailenson & Beall, 2006; Bailenson *et al.*, 2004) proposes that the possibilities that these real-time transformations raise can be classified into three categories or dimensions.

The first TSI dimension is *self-representation*. These transformations decouple the rendered appearance or behaviors of avatars from the actual appearance or behavior of the human driving it. That rendering can deviate from the actual state of the user. In a distance learning situation, it could be the case that some students learn better with teachers who utilize expressive gestures such as a smile, while some learn better with teachers with more stoic faces. In CMC, the teacher can be rendered differently to each student, with his facial gestures idiosyncratically depicted to each student in order to maximize that student's attention and learning style.

The second TSI dimension is *sensory abilities*. These transformations complement human perceptual abilities. One example is "invisible consultants," either algorithms or human avatars who can receive all sensory information from all interactants, but who are only visible (i.e., only rendered) to particular members of the CMC. These consultants can provide real-time summary information about the attentions and movements of other interactants (information which is automatically collected by tracking technology) or can scrutinize the actions of the user herself. For example, teachers using distance learning applications can utilize automatic registers that ensure that all students are looking in the direction of the teacher (a proxy for paying attention) to a sufficient degree.

The third TSI dimension is *situational context*. These transformations alter the spatial or temporal structure of a conversation. For example, the CMC can be optimally configured in terms of the geographical setup of a conference room. For example, every student in a class of 20 can sit directly in front of the virtual instructor, and perceive the rest of the students as sitting farther away. Furthermore, by altering the flow of rendered time in CMC, users can implement strategic uses of rewind and fast forward during a "real-time" interaction in a attempt to increase comprehension and efficiency.

Examples of TSI research

Here we review some previous findings relating to TSI, including published work as well as findings from some new work that has not yet been published.

Transforming the self

A majority of our work to date has centered upon examining transforming self- representation, largely because these are the types of transformation that are likely to occur across all types of CMC, compared to only media that involves very rich behavioral tracking and rendering such as immersive virtual reality.

Facial identity capture

Today, CMC involves the pervasive use of digital representations of people in video-conferences, static photographs accompanying emails and chats, as well as avatars used in online games. In a series of studies, we have demonstrated the effectiveness of algorithmic transformations that can be easily implemented in CMC which capitalize on human beings' disposition to prefer faces similar to their own (Bailenson, Garland, Iyengar, & Yee, 2006).

Similarity between two people instills altruism (Gaertner & Dovidio, 1977) and trust (DeBruine, 2002). Social explanations argue that people use physical similarity as a proxy for compatible interests and values (Zajonc *et al.*, 1987). Currently, political candidates tailor the information content of their mailings and televised messages to targeted demographic groupings (Iyengar *et al.*, 2001). Increasingly, they are in a position to vary salient attributes of their physical appearance, e.g. their weight, dress style, facial expression, or skin tone, depending on the audience in question. There is no reason to suspect that facial identity capture should be any different than clothing choice during digital campaigns.

In one study (Bailenson *et al.*, 2006), researchers passively acquired digital photographs of a national random sample of voting aged citizens. One week before the 2004 presidential election, participants completed a survey of their attitudes concerning George Bush and John Kerry while viewing photographs of both candidates side by side (see Figure 5.4). For a random one-third of the subjects, their own faces were morphed with Kerry while unfamiliar faces were morphed with Bush. For a different one-third, their own faces were morphed with Bush while unfamiliar faces were morphed with Kerry. The remaining one-third of the sample viewed unmorphed pictures of the candidates.

Postexperimental interviews demonstrated that not a single person detected that his or her image had been morphed into the photograph of the candidate. Participants were more likely to vote for the candidate morphed with their own face than the candidate morphed with an unfamiliar face. The use of facial identity capture was sufficient to change the outcome of the presidential election by a double-digit margin, according to a national random sample. In conclusion, using digital photographs, video images, and digital avatars allows people to dynamically morph representations during CMC. And by doing so, new, unique patterns of social influence will emerge.

Augmented gaze

Another TSI tool is *augmented gaze*: directing mutual gaze at more than a single interactant in a CMC system at once. Previous research has demonstrated that eye gaze is an extremely powerful tool for communicators seeking to garner attention, be persuasive, and instruct (see Segrin, 1993, for a review on this topic). People who use mutual gaze increase their ability to engage an audience as well as to accomplish a number of conversational goals.

In face-to-face interaction, gaze is zero-sum. In other words, if Person A

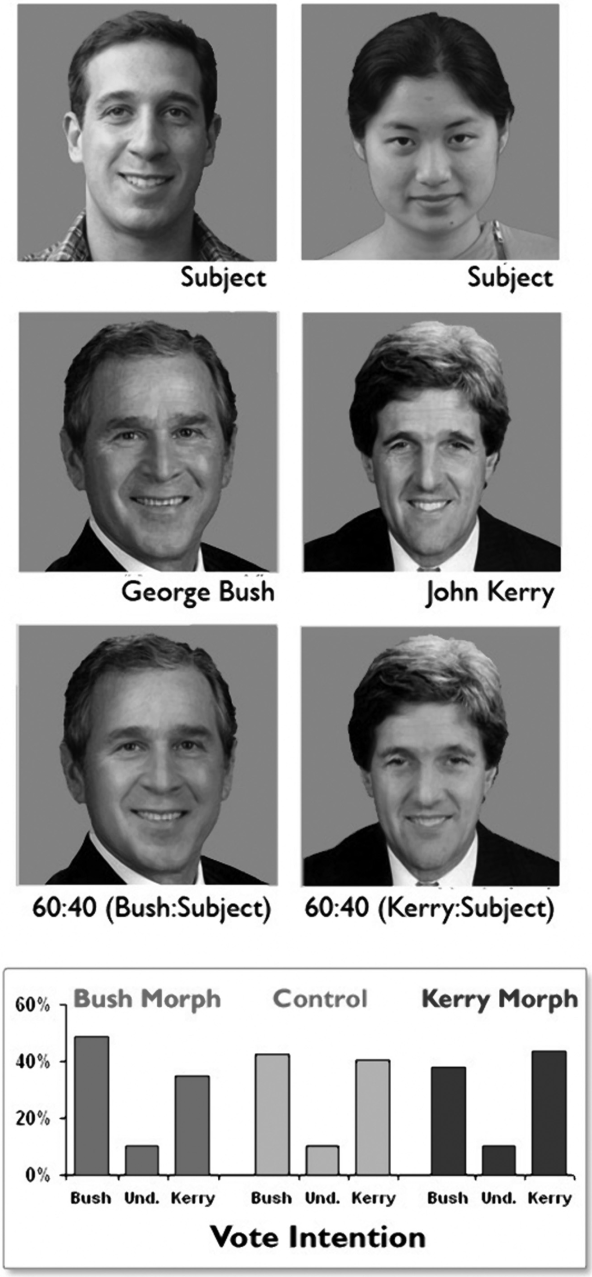

Figure 5.4 Two subjects, (top row), Bush and Kerry (2nd row) the morph of Subject 1 and Bush (3rd row left)), the morph of Subject 2 and Kerry (3rd row right), and the vote intention score by condition (bottom row). The difference in vote intention for Bush and Kerry by condition was significant ($p<.05$)

looks directly at Person B for 85 percent of the time, it is not possible for Person A to look directly at other people in the interaction for more than a total of 15 percent of the time. However, interaction among avatars during CMC is not bound by this constraint. The virtual environment as well as the other avatars in CMC is individually rendered for each interactant locally. As a result, Person A can have his avatar rendered differently for each other interactant, and appear to maintain mutual gaze with many interactants for the majority of the conversation, as Figure 5.5 demonstrates.

Augmented gaze allows interactants to perpetuate the illusion that they are looking directly at each person in an entire roomful of interactants. Three separate projects (Bailenson *et al*., 2004; Beall at al., 2003; Guadagno *et al*., 2005) have utilized a paradigm in which a single presenter read a passage to two listeners inside a collaborative virtual environment. All three interactants were of the same gender, wore stereoscopic, head-mounted displays, and had their head movements and mouth movements tracked and rendered, and the presenter's avatar either looked directly at each of the other two speakers simultaneously for 100 percent of the time (augmented gaze) or utilized normal, zero-sum gaze. Results across those three studies have produced and replicated three important findings: (1) participants never detected that the augmented gaze was not in fact backed by real gaze; (2) participants returned gaze to the presenter more often in the augmented condition than in the normal condition; and (3) participants (females to a greater extant than males) were more persuaded by a

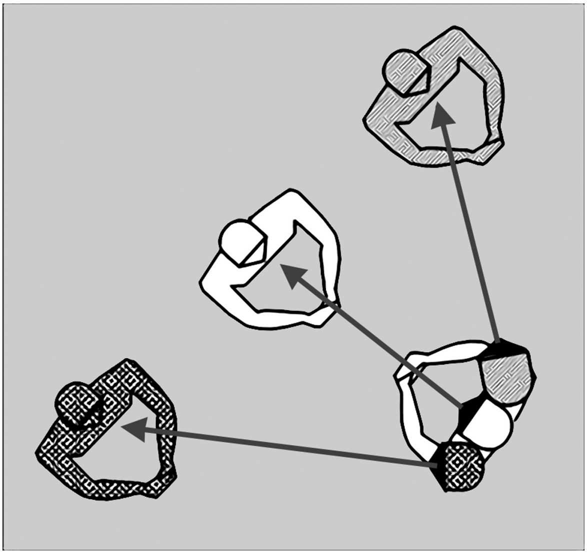

Figure 5.5 A schematic illustration of non-zero-sum gaze. Each interactant on the left perceives the speaker on the right gazing directly at him or her.

presenter implementing augmented gaze than a presenter implementing normal gaze. Augmented gaze will be a powerful tool in future computer-mediated communication. For applications such as distance learning, sales, online chatting and dating, utilizing computer guided gaze should have a high impact on learninsocial interaction.

The cyranoid

A *cyranoid* is an intermediary that communicates with a target person using the words or nonverbal behavior of another individual. Stanley Milgram described this concept and coined this term when he conducted a study in which participants interacted with an individual who, unbeknownst to them, was a cyranoid whose words were being controlled by a third party. In Milgram's words, cyranoids are: "People who do not speak thoughts originating in their own central nervous system: Rather, the words they speak originate in the mind of another person who transmits these words to the cyranoid by radio transmission." (Milgram *et al*., 1992, p. 337).

In one study (Guadagno *et al*., 2005), researchers examined whether a cyranoid (a virtual representation with verbal behavior controlled by one person and nonverbal behavior controlled by another) could be more persuasive than an avatar using augmented gaze (a virtual representation with verbal behavior controlled by one person and nonverbal behavior controlled by a computer algorithm). The cyranoid was instructed to engage or target a particular participant during the interaction by being persuasive with head movements. We expected that targeted nonverbal engagement by a third party (i.e., the cyranoid) would be more persuasive than a natural interaction, because a cyranoid can provide tailored nonverbal engagement without splitting attention between words and movements. Results indicated that as compared to control conditions, participants who were interacted with the cyranoid remembered more details of the persuasive passage, engaged in more mutual gaze with the presenter, liked the presenter better, and perceived more eye contact.

The Proteus effect

A great deal of social interaction occurs in virtual environments (Biocca & Levy, 1995; Parks & Floyd, 1995; Rheingold, 1993; Turkle, 1995; Walther, 1996; Walther *et al*., 1994; Yee, 2006), but the impact of our flexible self-representation within these environments has seldom been explored quantitatively. But given that social interaction in virtual environments revolves around a digital representation that can be altered in dramatic ways, it is important to understand how our altered self-representations—a process we term "the Proteus Effect." (Yee & Bailenson, 2006). We argue that just as men and women conform to gender roles, (i.e., *social role theory*, Eagly &

Wood, 1999) and just as the elderly conform to expected age stereotypes, (i.e., *self-stereotyping*, Levy, 1996), we might expect that people conform to stereotypical behaviors associated with their digital self-representations.

Two studies tested the Proteus Effect (Yee & Bailenson, 2006). The first explored the effect of attractiveness. Participants were immersed in a virtual room and saw their digital representation in a virtual mirror. Then they interacted with a confederate. In the *attractive* condition, participants were given an avatar of the same gender with an attractive face. In the *unattractive* condition, participants were given an avatar of the same gender with an unattractive face. These faces were chosen on the basis of a pretest and shown to differ significantly in terms of attractiveness ratings. Participants were then asked to perform an interpersonal distance task and a self-disclosure task. The results showed that participants in the attractive condition walked significantly closer and disclosed significantly more pieces of information than participants in the unattractive condition. In other words, the attractiveness of an avatar changes how friendly a person behaves towards other people in a virtual environment.

The second study explored the effect of height. The literature suggests that height is positively correlated with self-esteem (Judge & Cable, 2004). Thus, it was hypothesized that people given tall avatars would behave in a more confident way than those given short avatars, and three experimental conditions were developed. In the *tall* condition, participants had an avatar 15cm taller than the confederate's avatar. In the *short* condition, participants had an avatar 15cm shorter than the confederate's avatar. And in the *normal* condition, the participant's avatar was the same height as the confederate's avatar. These researchers employed a money-splitting negotiation task as a behavioral measure of confidence and found that participants in the tall condition were more willing to make unfair splits in their own favor, while participants in the short condition were more willing to accept unfair splits made by the confederate. Thus, this data again supported the Proteus Effect: Users given tall avatars became more confident than users given short avatars.

The Proteus Effect has broad implications for social interactions during CMC. We usually think of avatar creation as a one-way process, something of our own choosing, but the digital selves that we create in fact come to shape our behaviors in turn. Who we choose to be online changes how we behave.

Digital chameleons

Human behavioral researchers have long noted a synchronization and contagion of many verbal and nonverbal behaviors in social interactions, such as in speech patterns (Cappella & Panalp, 1981), posture (LaFrance, 1982), or mood (Neumann & Strack, 2000). More recently, researchers have found that automatic mimicry is a mechanism that increases social rapport in face-to-face interaction (Chartrand & Bargh, 1999; Chartrand & Jefferis, 2003). A subject who is mimicked by a confederate rates the confederate more positively after performing a task together, and subjects are more likely to mimic a confederate when there is a higher need for affiliation (Lakin & Chartrand, 2003). This line of evidence supports the claim that both unintentional (automatic) and intentional mimicry facilitates and expresses social affiliation and that the process is bi-directional—mimicry facilitates affiliation and prosocial behavior and affiliation goals increase mimicry. (Lakin *et al*., 2003).

CMC systems are uniquely suited for employing automatic mimicry for social advantage. Given that the system is already tracking a wide variety of actions and movements of interactants precisely, it becomes easy to mimic a person's movements accurately. It also becomes possible to build "nonverbal profiles" of users based on their past behaviors and save that into a database. And this mimicry might also be applied to embodied agents as an easy-to-implement algorithm for gaining social rapport with a human user.

To test this "digital chameleon" hypothesis, Bailenson and Yee (2005) conducted an experimental study in which undergraduate students were immersed in a virtual environment. In the virtual environment, participants were seated opposite an agent who presented an argument for approximately four minutes. The participant's head movements (i.e., pitch, yaw, and roll) were tracked by the VR system. In the mimic condition, the agent played back the participant's head movements with a four-second delay. In the recorded condition, the agent played back the recording of a different participant from the mimic condition.

The results from the study showed that participants in the mimic condition were more likely to pay attention to the agent in terms of gaze and agree with the agent's argument than participants in the recorded condition. More importantly, less than 5 percent of the participants had detected the mimicry in the post-experiment questionnaire. These findings have substantial implications. Given the precision with which CMC systems can track an individual's nonverbal behavior, it allows avatars and agents to use automatic mimicry for social advantage. These findings also show that such an algorithm is easy to implement, requiring no preexisting library or syntax of nonverbal gestures to function. Thus, the interaction and the meaning of specific nonverbal gestures do not even need to be understood by the system for this transformation to be effective.

Transforming sensory abilities

Government funding agencies issued a major push in the late 1990s with a research agenda called Augmented Cognition (see Schmorrow & Kruse, 2004 for a detailed history), designing computer interfaces to extend the limitations of normal human cognition. One major rationale for this work was to provide digital wearable displays that could increase the working memory of people by allowing them to be able store cognitive information on displays as opposed to having to keep them actively stored in memory. Similarly, we have been conducting CMC research to provide augmentations of social sensory abilities. These transformations compliment human abilities to draw inferences about the social world. In this section we discuss two examples of such augmentations.

Multilateral perspective taking

Many CMC systems, such as online games or video-conferences, use multiple viewpoints or virtual cameras to allow users to decouple their visual point of view from that of their avatars (i.e., the normal view from the eyes). In theory, in any digital communication system, it should be possible for an interactant to take a visual point of view from any single point in the virtual room. In other words, it is possible for Person B to disconnect the area of perception from the area in which Person A perceives her. Figure 5.6 illustrates this transformation.

In Figure 5.6, Person B is implementing a *multilateral perspective*. Specifically, she is choosing to adopt the sensory perspective of Person A during the conversation. In other words, she has left her own point of view and become a passenger to Person A, by viewing a digital world that is not contingent on her own movements, but instead a digital world that is

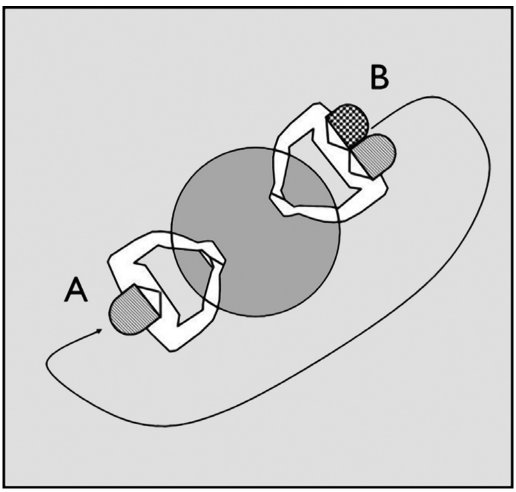

Figure 5.6 Person B takes on multilateral perspectives: she can experience the CVE her own perspective and the perspective of Person A at the same time.

contingent on Person A's movements. As a result, she sees herself in real time from behind the eyes of her conversational partner. Either by shifting her entire field of view to the spatial location of other avatars in the interaction, or by popping up "field of view windows" in corners of the virtual display, an interactant can unobtrusively occupy the home space of any avatar in the CVE.

Research (Gehlbach *et al.*, 2005) is examining multilateral perspectives in a negotiation scenario inside a CVE. Previous work has used either role playing (Davis *et al.*, 1996) or observational seating arrangements (Taylor & Fiske, 1975) to cause subjects to take on the perspectives of others in a conversation, demonstrating more efficient and effective interactions. Equipping an interactant with the real-time ability to see one's avatar from another point of view should enhance these effects. In our work in progress, we are predicting more cooperative solutions in simulations in which negotiators can occupy the field of view of their opponents.

Behavioral flags

During any interaction, meaningful events occur that involve complex behaviors, verbal and nonverbal utterances. In order to render the actions of participants to one another in a CMC, it is necessary to capture all information about those actions. The current study examines how interactants benefit from receiving real-time, summary information about the social actions of themselves and others. We are planning to examine one-on-one scenarios such as tutoring, negotiation, and sales pitches, as well as one-on-many scenarios such as class lectures. Figure 5.7 illustrates a CMC system that displays information flags over the heads of three users. In these instances, one or more of the interactants in an immersive virtual environment CMC system will receive real-time information about the following behaviors:

1 *Nod/Head Shake Detection*. Using a simple device that tracks head orientation, it is possible to detect agreement nods or disagreement shakes using spectral analysis on the head orientation data. The ratio of these behaviors for a given user should be indicative of agreement and comprehension.
2 *Facial Expressions*. Using advanced software by Nevenvision which uses computer vision to automatically track facial features in real time (approximately 10 hz), we have developed and tested a system that detects simple facial expressions, such as smiles and frowns. We will continue to attempt to isolate additional expressions.
3 *Gaze Behavior*. In previous work (Beall *et al.*, 2003), we have used head tracking equipment to determine when people look in each other's eyes (i.e. mutual gaze). This tool will help a user know how often he

has looked at all the other interactants as well as how often they have looked at him.

4 *Speaking Frequency*. Using a simple microphone that records the frequency of speech, we have previously automatically computed the percentage of time each person is speaking (Bailenson *et al.*, 2002).

Transforming situational context

This is the dimension of our theoretical framework which has received the least empirical attention as of yet. While technological development and empirical design is underway for multiple studies transforming a user's context, we discuss only one of these ongoing research studies in the current section, called *Transformed Conformity*.

Conformity is one of the most powerful forms of social influence (Asch, 1955; Festinger, 1954). Previous research in collaborative virtual environments (Blascovich *et al.*, 2002; Swinth & Blascovich, 2002) has demonstrated that participants conform to the behaviors of other people in immersive virtual reality, regardless of whether they are avatars (representations controlled by other people) or agents (representations controlled by the computer). In current work, we are examining the effect of overriding the behaviors of other group members in CMC. In other words, for

Figure 5.7 Three participants with the behavioral flags translucently displayed over their heads. Only the presenter (behind the podium) in this CVE can see the behavioral flags.

any given participant, instead of seeing the actual behaviors of his or her peers, participants can see transformed behaviors. The goal of this work is to examine the effectiveness of presenters who create a specific type of audience via transformed conformity.

We are currently designing and running studies in which participants are present in the same collaborative virtual environment, and manipulating the *types of transformed behaviors* that each participant perceives of his or her neighbors. Each participant either sees the actual behaviors (e.g., facial expressions, direction of eye gaze, nodding and head shaking behaviors) of the other group members or sees transformed behaviors that are created to induce participants to conform to a certain standard. For example, positive attention behaviors include overriding actual behaviors to make the surrounding students look at the presenter, nod, smile, and ask questions. On the other hand, negative attention behaviors include frowning, demonstrating boredom expressions, sitting with eyes closed and gaze aversion. In pilot studies, subjects learning in positive learning environments resulting from transformed conformity are demonstrating more learning, persuasion, and mutual gaze than subjects in control conditions.

TSI and social influence theory

In the previous section, we described a number of studies showing that TSI can be used for people to achieve social influence. Indeed, nearly every single example from above features interactants transforming their avatars, senses, or context in order to strategically accomplish some goal relating to teaching or persuasion (e.g., facial identity capture, digital chameleons, etc.). In order to provide a theoretical framework to guide research in TSI, we look to the model most relevant to this work. Blascovich and colleagues (Blascovich, 2002; Blascovich *et al.*, 2002) proposed a model of social influence during CMC in virtual environments. This model was primarily developed to understand social interaction inside immersive virtual environments, but it applies equally well to other types of CMC.

As Figure 5.8 depicts, there is a tradeoff between realism (the degree to which human representations look and behave as they would in the physical world) on the vertical axis and perceived agency (the extent to which the interactant thinks they are interacting with another actual human being) on the horizontal axis. The higher the realism, particularly communicative realism (e.g., facial expressions), then the less perceived agency needed to achieve social influence and vice versa. Hence, according to the model, social influence is likely to occur when either realism or agency are high, or both.

According to this model, individuals consciously respond differently to virtual representations that are computer controlled (agents) than they will to human-controlled virtual representations (avatars) at all but the highest levels of realism, as the threshold of social influence demonstrates in Figure 5.8. Specifically, embodied agents need to display more behavioral realism than avatars in order for conscious social influence to take place. However, the model specifies that the agent-avatar distinction is less important for unconsciously controlled low-level reflexive or automatic behaviors (e.g., maintaining appropriate interpersonal distance; facial mimicry).

According to Blascovich *et al.* (2002), the realism variable in the model is regarded as a latent variable, which can only be assessed or manipulated via manifest realism variables. The model specifies the latter in a hierarchical fashion such that communicative (i.e., social) realism is the most important manifest variable. Communicative realism involves movements (e.g., vocal chords to produce sounds; facial muscle and gestural muscle movements to produce nonverbal signals). Anthropometric realism (e.g., the shape or morphology of the virtual human representation) is important in its service of communicative realism (e.g., one cannot have a hand gesture or lip movements without an arm/hand or face, respectively). Photographic realism is less important dynamically but can be important in terms of social or group identity.

This model provides a general framework to make predictions of TSI manipulations as well as to interpret results, and suggests that photographic realism is much less important than realism associated with behaviors particularly communicative ones such as facial expressions, gestures, head movements, etc. Furthermore, it suggests that in terms of low-level or automatic behaviors, there should be no differences in terms of perceived agency. That is, both known agents and avatars should have the same effects on such unconsciously generated behaviors. For example, either

Figure 5.8 A model of social influence in CVEs

should be able to elicit a startle, defensive, or orientation response on the part of a user.

In the empirical section, we have discussed a number of manipulations in experiments which allow someone to achieve higher amounts of social influence via TSI. The question becomes, how do we apply these findings to a model of social influence? The answer is complicated: On the one hand, TSI should decrease realism, because the actual avatar projected is actually different from the person behind the avatar, and consequently less realistic objectively. On the other hand, the perception of the transformed avatar for the audience may in fact be perceived as more realistic because that avatar uses an optimal set of cues to achieve some conversational goal.

Walther's (1992; 1996) Social Information Processing Theory may provide guidance for this application. Walther argues that, in some instances, CMC can actually be "hyperpersonal," or more intimate than face-to-face settings, due to the fact that in CMC one can project an ideal self and redirect cognitive resources that would usually be applied to nonverbal behavior. Applying TSI may allow a user to become "hyper-realistic"—while the avatar is different from the actual self, it is idealized to become more real than would be possible in face-to-face settings. Of course, this is conjecture at this point, but future work should empirically examine the relationship between realism, hyperpersonal perceptions, TSI and social influence to shed light on these theoretical relations.

Ethics and implications

In sum, when people enter into new and novel types of CMC such as immersive virtual environments, some expectation of nonveridical rendering of others' behavior is most likely inevitable. However, when viewing more traditional types of CMC, such as two-dimensional video feeds, images on web sites, voices enhanced by digital algorithms on cell phones, other players in online video games and text in chat rooms, we may not be so rigorous in our skepticism concerning the authenticity of form and behavior. The potential for using TSI for abuse in all forms of digital communication certainly warrants attention.

There is an underlying Orwellian theme behind TSI strategies such as identity capture, augmented gaze, and digital mimicry. Some might argue that these tools would be better left out of the hands of advertisers, politicians, and anyone else who may seek to influence people. After all, TSI strategies allow them to gain advantages in persuasion and even in voting decisions that they would not have otherwise. On the other hand, "manipulative" strategies are nothing new to politicians. From sporting a sudden tan to selecting which video clips or photos to send to constituents, politicians have a great deal of control over how they present themselves to the public via different communication channels. And if it is inevitable that TSI strategies will be employed in the near future, then perhaps the most important thing is make people aware of these manipulations–much as how people now widely assume that magazine cover models have been airbrushed.

Once TSI strategies become widely known however, another possible scenario might occur. People may begin to distrust interaction that occurs in virtual spaces. For example, we could imagine scenarios where the premium package from your local internet service provider is not related to access speeds, but access to sophisticated TSI suites. Widespread use of TSI might lead to an infinite regression, a complete distrust in the medium itself. However, this line of argument fails to take into account the prevalence and acceptance of nondigital TSIs in our everyday lives (see Figure 5.2). The claim that digital TSIs will cause distrust assumes that people want to see and interact with each other without any intentional alterations. A cursory glance at modern societies reveals otherwise. One main function of clothing is to conceal the naked body; deodorants (and the sheer diversity of bath and shower products) are used to suppress our natural scents. If anything, there are certain TSIs that our society demands that we perform.

There is also an assortment of nondigital TSIs that are not socially mandatory, and these typically fulfill a cosmetic role. For example, these include: hair coloring, teeth whitening, haircuts, and make-up. In most of these cases, instead of shunning a person for deliberately deceiving others, we in fact typically compliment them on the improvement in their appearance. This is also the case with weight loss and dieting programs. We do not distrust a person because they have "deceptively" tried to create a new appearance, but instead accept their new appearance as an improvement. Overall, it appears that alterations that improve one's social presentation are in fact encouraged by society.

Of course, like any new technology, it takes time for a culture to develop norms for the technology's use. As CMC becomes more advanced and prevalent, it will be fascinating to monitor the progress of TSI strategies as well as technology designed to detect and foil the nonveridical rendering of appearance and behaviors. In the meantime, TSI in CMC present spectacular opportunities for social scientists studying communication and social interaction.

References

Asch, S. (1955) "Opinions and Social Pressure." *Scientific American, 193*, 5.

Bailenson, J. N. (2006) "Transformed social interaction in collaborative virtual environments." In P. Messaris. and L. Humphreys, (eds.), *Digital Media: Transformations in Human Communication*. New York: Peter Lang, pp. 255–264.

Bailenson, J. N., & Beall, A. C. (2006) "Transformed social interaction: Exploring the digital plasticity of avatars." In R. Schroeder, & A. Axelsson, (eds.), *Avatars at Work and Play: Collaboration and Interaction in Shared Virtual Environments*, Springer-Verlag, pp. 1–16.

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

Bailenson, J., & Blascovich, J. (2004) "Avatars." In W. S. Bainbridge (ed.), *Encyclopedia of Human-Computer Interaction*. Great Barrington, MA: Berkshire Publishing Group.

Bailenson, J. N., & Yee, N. (2005) "Digital Chameleons: Automatic assimilation of nonverbal gestures in immersive virtual environments." *Psychological Science, 16*, 814–819.

Bailenson, J. N., Iyengar, S., & Yee, N. (2006) Facial Identity capture in the 2004 presidential election. Manuscript submitted for publication.

Bailenson, J.N., Garland, P., Iyengar, S., & Yee, N. (2006) "Transformed Facial Similarity as a Political Cue: A Preliminary Investigation." *Political Psychology, 27*, 373–386.

Bailenson, J., Beall, A., 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.

Beall, A., Bailenson, J., Loomis, J. Blascovich, J., & Rex, C. (2003) "Non-Zero-Sum Mutual Gaze in Collaborative Virtual Environments." Paper presented at the Proceedings of HCI International, Crete.

Biocca, F., & Levy M. R. (1995) *Communication in the Age of Virtual Reality*. Hillsdale, NJ, L. Erlbaum Associates.

Blascovich, J. (2002) "Social influence within immersive virtual environments." In R. Schroeder (ed.), *The Social Life of Avatars*. London: Springer-Verlag, pp. 127–145.

Blascovich, J., Loomis, J., Beall, A., Swinth, K., Hoyt, C., & Bailenson, J. (2002) "Immersive virtual environment technology as a methodological tool for social psychology." *Psychological Inquiry, 13*(2), 103–124.

Burgoon, J., Bonito, J., Bengtsson, B., Ramirez, A., Dunbar, N., & Miczo, N. (2000) "Testing the interactivity model: Communication processes, partner assessments, and the quality of collaborative work." *Journal of Management Information Systems, 16*, 33–56.

Cappella, J., & Panalp, A. (1981) "Talk and silence sequences in informal conversations. Interspeaker influence." *Human Communication Research, 7*, 117–132.

Chartrand, T. L., & Bargh, J. A. (1999) "The chameleon effect: The perception-behavior link and social interaction." *Journal of Personality & Social Psychology, 76*(6), 893–910l.

Chartrand, T. L., & Jefferis, V. E. (2003) "Consequences of automatic goal pursuit and the case of nonconscious mimicry." In J. P. Forgas & K. D. Williams (eds.), *Social Judgments: Implicit and Explicit Processes*. New York: Cambridge University Press, pp. 290–305.

Davis, M., Conklin, L., Smith, A., & Luce, C. (1996) "Effect of perspective taking on the cognitive representations of persons: A merging of self and other." *Journal of Personality and Social Psychology, 70*, 713–726.

DeBruine, L. M. (2002) "Facial resemblance enhances trust." *Proceedings of the Royal Society of London B, 269*, 1307–1312.

Dion, K., Berscheid, E., & Walster, E. (1972) "What is beautiful is good." *Journal of Personality and Social Psychology, 24*, 285–290.

Eagly, A. H., & Wood, W. (1999) "The origins of sex differences in human behavior: Evolved dispositions versus social roles." *American Psychologist, 54*, 408–423.

ESA (2005) *Entertainment Software Association: Facts and Research*, from http://www.theesa.com/facts/gamer_data.php.

Festinger, L. (1954) "A theory of social comparison processes." *Human Relations, 7*, 117–140.

Gaertner, S. L., & Dovidio, J. F. (1977) "The subtlety of white racism, arousal and helping behavior." *Journal of Personality and Social Psychology, 35*, 691–707.

Gehlbach, H., Bailenson, J., Yee, N., & Beall, A. (2005) Perspective taking and negotiation in collaborative virtual environments. Unpublished manuscript.

Giles, H., & Claire, R. (1979) *Language and Social Psychology*. Oxford: Blackwell.
Goffman, E. (1959) *The Presentation of Self in Everyday Life*. New York: Anchor Books.
Guadagno, R. E., Bailenson, J. N., Beall, A. C., Dimov, A., & Blascovich, J. (2005) "Transformed social interaction and the cyranoid: The impact of non-verbal behavior on persuasion in an immersive virtual environment." urly quotes Unpublished manuscript, University of California, Santa Barbara.
Higham, P. A., & Carment, W. D. (1992) "The rise and fall of politicians: The judged heights of Broadbent, Mulroney and Turner before and after the 1988 Canadian federal election." *Canadian Journal of Behavioral Science, 24*, 404–409.
Iyengar, S., Lowenstein, D. L., & Masket, S. (2001) "The stealth campaign: Experimental studies of slate mail in California." *The Journal of Law and Politics, 17*, 295–232.
Johnston, L. (2002) "Behavioral mimicry and stigmatization." *Social Cognition, 20*(1), 18–35.
Joinson, A. N. (2001) "Self-disclosure in computer-mediated communication: The role of self-awareness and visual anonymity." *European Journal of Social Psychology, 31*, 177–192.
Jones, E. E. (1964). *Ingratiation*. New York: Irvington.
Jones, E. E., & Pitman, T. S. (1982) "Toward a general theory of strategic self-presentation." In J. Suls (ed.), *Psychological perspectives on the Self*, Vol. 1. Hlllsdale, NJ: General Learning Press, pp. 231–262.
Judge, T. A., & Cable, D. M. (2004) "The effect of physical height on workplace success and income: Preliminary test of a theoretical model." *Journal of Applied Psychology, 89*, 428–441.
Konijn, E. A., & Hoorn, J. F. (2004) "Reality-based genre preferences do not direct personal involvement." *Discourse Processes, 38*, 219–246.
LaFrance, M. (1982) "Posture mirroring and rapport." In M. Davis (ed.), *Interaction Rhythms: Periodicity in Communicative Behavior*, New York: Human Sciences Press, pp. 279–298.
Lakin, J. L., & Chartrand, T. L. (2003). Using nonconscious behavioral mimicry to create affiliation and rapport." *Psychological Science, 14*(4), 334–339.
Lakin, J. L., Jefferis, V. E., Cheng, C. M., & Chartrand, T. L. (2003) "The chameleon effect as social glue: Evidence for the evolutionary significance of nonconscious mimicry." *Journal of Nonverbal Behavior, 27*(3), 145–162.
Langlois, J., Kalakanis, L., Rubenstein, A., Larson, A., Hallam, M., & Smoot, M. (2000) "Maxims or myths of beauty?: A meta-analytic and theoretical review." *Psychological Bulletin, 126*, 390–423.
Levy, B. (1996) "Improving memory in old age through implicit self-stereotyping." *Journal of Personality and Social Psychology, 71*, 1092–1107.
Milgram, S., Sabini, J., & Silver, M. (1992) *The Individual in a Social World: Essays and Experiments*. New York: McGraw-Hill.
Nass, C., & Brave, S. (2005) *Wired for Speech: How Voice Activates and Advances the Human-Computer Relationship*. Cambridge, MA: MIT Press.
Neumann, R., & Strack, F. (2000) "Mood contagion: The automatic transfer of mood between persons." *Journal of Personality & Social Psychology, 79*, 211–223.
Parks, M. R., & Floyd, K. (1995) "Friends in cyberspace: Exploring personal relationships formed through the Internet." Paper presented at the International Communication Association, Albuquerque, NM.
Project, P. I. a. A. L. (2004) "*53 Million in US use IM*," from http://www.emarketer.com/Article.aspx?1003031.
Provine, R. (1986) "Yawning as a stereotyped action pattern and releasing stimulus." *Ethology, 72*, 109–122.
Provine, R. (1992) "Contagious Laughter: Laughter is sufficient stimulus for laughs and smiles." *Bulletin of the Psychonomic Society, 30*, 1–4.
Reidsma, D., Akker, R., Rienks, R., Poppe, R., Poppe, R., Nijholt, A., Heylen, D., & Zwiers, J.

(2005) "Virtual meeting rooms: from observation to simulation." In *Proceedings of Social Intelligence Design, 2005*.

Rheingold, H. (1993) *The Virtual Community: Homesteading on the Electronic Frontier*. Reading, MA: Addison-Wesley.

Schlenker, B. R. (1980) *Impression Management: the Self-Concept, Social Identity, and Interpersonal Relationships*. Monterey, CA: Brooks/Cole.

Schroeder, R. (ed.), (2002) *The Social Life of Avatars*. London: Springer-Verlag.

Schmorrow, D. D., & Kruse, A. A. (2004) "Augmented Cognition." In W. S. Bainbridge (ed.), *Berkshire Encyclopedia of Human-Computer Interaction*. Great Barrington, MA: Berkshire Publishing Group, pp. 54–59.

Segrin, C. (1993) "The effects of nonverbal behavior on outcomes of compliance gaining attempts." *Communication Studies, 44*, 169–187.

Sproull, L., & Kiesler, S. (1986) "Reducing social context cues: Electronic mail in organizational communication." *Management Science, 32*, 1492–1512.

Sundar, S. S., Kalyanaraman, S., & Brown, J. (2003) "Explicating website interactivity: Impression-formation effects in political campaign sites." *Communication Research, 30*, 30–59.

Swinth, K., & Blascovich, J. (2002) "Perceiving and responding to others: Human-human and human-computer social interaction in collaborative virtual environments." Paper presented at the Proceedings of the 5th Annual International Workshop on PRESENCE. Porto, Portugal.

Taylor, S., & Fiske, S. (1975) "Point of view and perception so causality." *Journal of Personality and Social Psychology, 32*, 439–445.

Turk, D. J., Heatherton, T. F., Kelley, W. M., Funnell, M. G., Gazzaniga, M. S., & Macrae, C. N. (2002) "Mike or me? Self recognition in a split-brain patient." *Nature Neuroscience, 5*, 841–842.

Turkle, S. (1995) *Life on the Screen: Identity in the Age of the Internet*. New York: Simon and Schuster.

Walther, J. (1992) "Interpersonal Effects in Computer-Mediated Interaction: A Relational Perspective," *Communication Research, 19*, pp. 52–90.

Walther, J. B. (1996) "Computer-mediated communication: Impersonal, interpersonal, and hyperpersonal interaction." *Communication Research, 23*(1), 3–43.

Walther, J. B., Anderson, J. F., & Park, D. (1994) "Interpersonal effects in computer-mediated interaction: A meta-analysis of social and anti-social communication." *Communication Research, 21*, 460–487.

Woodcock, B. (2005) *MMOG Chart*, from http://www.mmogchart.com/

Yee, N. (2006, in press) "The Demographics, Motivations, and derived experiences of users of massively multi-user online graphical environments." *Presence: Teleoperators and Virtual Environments*.

Yee, N., & Bailenson, J. N. (2007) "The Proteus effect: Self Transformations in Virtual Reality." *Human Communication Research*, *33*, 271–290.

Young, T. J., & French, L. A. (1996) "Height and perceived competence of U. S. Presidents." *Perceptual and Motor Skills, 82*, 1002.

Zajonc, R. B., Adelmann, P. K., Murphy, S. T., & Niendenthal, P. M. (1987) "Convergence in the physical appearance of spouses." *Motivation and Emotion, 11*, 335–346.