Transformed Social Interaction:

Exploring the Digital Plasticity of Avatars

Jeremy N. Bailenson

Department of Communication

Stanford University

Andrew C. Beall

Department of Psychology

University of California, Santa Barbara

Bailenson, J.N. & Beall, A.C. (2005, in press). Transformed Social Interaction: Exploring the Digital Plasticity of Avatars. In Schroeder, R. & Axelsson, A.'s (Eds.), *Avatars at Work and Play: Collaboration and Interaction in Shared Virtual Environments*, Springer-Verlag.

<Ch1> Transformed Social Interaction: Exploring the Digital Plasticity of Avatars

<h1>Introduction

What does it mean to be you? How drastically can a person change and still remain, in the eyes of either themselves or their peers, the same person? Until recently, these questions were typically asked in the context of philosophy, psychoanalysis, or science fiction. However, the increasingly common use of avatars during computer-mediated communication, collaborative virtual environments (CVEs) in particular, are quickly changing these once abstract questions into practical quandaries that are fascinating, thought-provoking, potentially paradigm shifting for those who study social interaction, and potentially devastating to the traditional concept of human communication.

Historically, even before the advent of computers, people have demonstrated a consistent practice of extending their identities. As Turkle [1] points out:

> The computer of course, is not unique as an extension of self. At each point in our lives, we seek to project ourselves into the world. The youngest child will eagerly pick up crayons and modeling clay. We paint, we work, we keep journals, we start companies, we build things that express the diversity of our personal and intellectual sensibilities. Yet the computer offers us new opportunities as a medium that embodies our ideas and expresses our diversity (p. 31).

Extending one's sense of self in the form of abstract representation is one of our most fundamental expressions of humanity.

But abstract extension is not the only manner in which we manipulate the conception of the self. In addition to using abstract means to extend one's identity, humans also engage in the practice of using tangible means to transform the self. Table 1 demonstrates some of these self transformations that occur currently, without the use of digital technology. Before the dawn of avatars and computer-mediated communication, this process of self transformation was minor, incremental, and required vast amounts of resources.

However, given the advent of collaborative virtual reality technology [2-5], as well as the surging popularity of interacting with digital representations via collaborative desktop technology [6], researchers have begun to systematically explore this phenomenon of *Transformed Social Interaction* [7]. TSI involves novel techniques that permit changing the nature of social interaction by providing interactants with methods to enhance or degrade interpersonal communication. TSI allows interactants themselves, or alternatively a moderator of the CVE, to selectively filter and augment the appearance, verbal behavior, and nonverbal behavior of their avatars. Furthermore, TSI also allows the interactants to filter the context in which an interaction occurs. In our previous work outlining the theoretical framework of TSI, we provided three dimensions for transformations during interaction.

The first dimension of TSI is transforming *sensory abilities*. These transformations augment human perceptual abilities. For example, one can have 'invisible consultants' present in a collaborative virtual environment, ranging from other avatars of assistants rendered only to you who scrutinize other interactants, to algorithms that give you real-time summary statistics about the movements and attentions of others (which are automatically collected in a CVE in order to render behaviors). As a potential application, teachers using distance learning applications can have “attention monitors” that automatically use eye gaze, facial expressions and other gestures as a mechanism to localize students who may not understand a given lesson. That teacher can then tailor his or her attention more towards the students higher in need. As another example, teachers can render virtual nametags (displayed to the teacher only) inserted over their students’ avatars. Consequently, even in a distance learning classroom of hundreds, the students’ names will always be at an instructor’s disposal without having to consult a seating chart or a list.

The second dimension is *situational context*. These transformations involve changes to the temporal or spatial structure of an interaction. For example, each interactant can optimally adjust the geographical configuration of the room—in a distance learning paradigm, every single student in a class of twenty can sit right up front, next to the teacher, and perceive his or her peers as sitting behind. Furthermore, real-time use of “pause” and “rewind” during an interaction (while one’s avatar exhibits stock behaviors produced by an “auto-pilot” algorithm) may be quite an effective tool to increase comprehension and productivity during interaction. Another example of transforming the situational contexts is to utilize *multilateral perspectives*. In a normal conversation, interactants can only take on a single perspective—their own. However, in a CVE, one can adopt the visual point of view of any avatar in the entire room. Either by bouncing her entire field of view to the spatial location of other avatars in the interaction, or by keeping “windows” in the corners of the virtual display that show in real-time the fields of views of other interactants, it is possible for an interactant to see the behavior of her own avatar, as they occur, from the eyes of other interactants. Previous research has used either role playing scenarios [8] or observational seating arrangements [9] to cause experimental subjects to take on the perspectives of others in an interaction, and has demonstrated that this process is an extremely useful tool for fostering more efficient and effective interactions. Equipping an interactant with the real-time ability to see one’s avatar from another point of view should only enhance these previous findings concerning the benefits of taking other perspectives.

The third dimension of TSI is *self representation*. These transformations involve decoupling the rendered appearance or behaviors of avatars from the human driving the avatar. In other words, interactants choose the way in which their avatars are rendered to others in the CVE, and that rendering can follow as closely or as disparately to the actual state of the humans driving the avatars as they so desire. The focus of this paper will be to discuss this third dimension in greater detail. While transforming situational contexts and sensory abilities are fascinating constructs, thoroughly discussing all three dimensions is beyond the scope of the current work.

This idea of decoupling representation from actual behavior has received some attention from researchers previously exploring CVEs. For example, [10] as well as [11] discussed *truthfulness* in representation, Biocca [12] introduced a concept known as *hyperpresence*, using novel visual dimensions to express otherwise abstract emotions or behaviors, and, moreover, numerous scholars debate the pros and cons of abstract digital identities [1, 13]. Furthermore, Jaron Lanier, considered by many to be one of the central figures in the history of immersive virtual reality, often makes an analogy between the human using immersive virtual reality and the “aplysia”, a sea-slug that can quickly change its surface features such as body shape and skin color. Before virtual reality, humans had to resort to makeup, plastic surgery, or elaborate costumes to achieve these goals. William Gibson [14] may have put it best when he declared that, once the technology supports such transformations, it is inevitable that people take advantage of “the infinite plasticity of the digital,” (pg. 117).

In sum, the idea of changing the appearance and behaviors of one’s representation in immersive virtual reality has been a consistent theme in the development of the technology. The goals of the Transformed Social Interaction paradigm are threefold: 1) to explore and actually implement these strategies in collaborative virtual environments, 2) to put human avatars in CVEs and to measure which types of TSI tools they actually use during interaction, and 3) to examine the impact that TSI has on the effectiveness of interaction in general, as well as the impact on the specific goals of particular interactants. In the current paper, we provide an overview of the empirical research conducted to date using avatars to examine TSI, and then discuss some of the broader implications of these digital transformations.

<h1> Transforming Avatar Appearance

This section reviews a series of TSI applications concerning the static appearance of one’s avatar, some of which have been already tested using behavioral science studies in CVEs, others that have yet to receive empirical examination.

<h2> Identity Capture

The nature of a three-dimensional model used to render an avatar lends itself quite easily to applying known algorithms that transform facial structure according to known landmark points on the head and face. Once a face is digitized, there are an infinite number of simple morphing techniques that alter the three-dimensional structure and surface features of that face. This practice can be a powerful tool during interaction.

For example, persuaders can absorb aspects of an audience member's identity to create implicit feelings of similarity. Imagine the hypothetical case in which Gray Davis (the past governor of California, depicted in the leftmost panel of Figure 2) is attempting to woo the constituents of a locale in which the voters are primarily fans of Arnold Schwarzenegger (the governor of California that ousted Davis) depicted in the rightmost panel of Figure 2.

Research in social psychology has demonstrated large effects of similarity on social influence, in that a potential influencer who is more similar to a given person (compared to a less similar influencer) is considered more attractive [15] and persuasive [16], is more likely to make a sale [17], and is more likely to receive altruistic help in a dire situation [18]. Consequently, using digital technology to "absorb" physical aspects of other interactants in a CVE may provide distinct advantage for individuals who seek to influence others, either in a positive manner (e.g., a teacher during distance learning), or in a manner not so wholesome (e.g., a politician trying to underhandedly co-opt votes). Moreover, this type of a transformation may be particularly effective in situations in which the transformation remains implicit [19]. In other words, the effect of the transformation may be strongest when CVE interactants do not consciously detect their own face morphed into the face of the potential influencer.

To test this hypothesis, we brought Stanford University undergraduate students into the lab and used a simple morphing procedure with MagicMorph software [20, 21] to blend their faces in with an unfamiliar politician, Jim Hahn, a mayor of Los Angeles. Figure 4 depicts images of two undergraduate students as well as two blends that are each compromised of 60 percent of Jim Hahn and 40 percent of their own features. The main hypothesis in this study [22], was that participants would be more likely to vote for a candidate that is morphed with their own face than a candidate that is morphed with someone else's face. In other words, by capturing a substantial portion of a voter's facial structure, a candidate breeds a feeling of familiarity, which is an extremely effective strategy for swaying preference [23].

Our findings in this study demonstrated two important patterns. First, out 36 participants, only two detected that their own face was morphed into the candidate, even when we explicitly asked them to name one person like whom the candidate looked. Interestingly, their responses often demonstrated an implicit similarity (e.g., "He looks like my grandfather," or "He looks really familiar but I am not sure who he is,"), but very rarely indicated a detection of the self. Second, overall there was a preference for candidates that were morphed with the self over candidates that were morphed with others, though the effect was strongest for white male participants (who were similar enough to the picture of Jim Hahn to create a successful morph) and for people interested in politics (who ostensibly were more motivated to pay attention to the photograph of the candidate). In sum, very few participants noticed that their face was morphed into the political candidate, but implicitly the presence of themselves in the candidate gave the candidate a greater ability to influence those participants.

<h2> Team Face

A related study [24] examined the use of TSI for collaborative teams by creating a “Team Face”. Given the underlying notion that teams function more cooperatively when they embrace commonalties (e.g., dress codes, uniforms) it is logical to consider that organizations would consider extending these team features to the rendering of avatars. Consider the faces in Figure 4. The face on the far right is a morphed avatar that includes the faces from all four of the participants at equal contributions. In our study, participants (32 in total: four sets of four participants of each gender) received two persuasive messages: one delivered by their own team face, and one delivered by a team face that did not include their own face.

In this study, only three participants noticed their own face present inside the team face when explicitly asked to name one person like whom the face looked. In regards to persuasion, our results indicated that when participants received a persuasive message from an avatar wearing the team face, they were more likely to scrutinize the arguments. Specifically, arguments that were strong (determined by pre-testing) were seen as stronger when received by one’s own team face than when received by a different team face, and the opposite pattern occurred for weak arguments.

This pattern is quite consistent with what would be predicted by the elaboration-likelihood model of Petty and Cacioppo [25]. According to that model, people processing a persuasive message utilize either the central route (i.e., dedicate cognitive resources towards actually working through the logical strengths and weaknesses of an argument) or the peripheral route (i.e., analyze the message only in terms of quick heuristics and surface features). In the study using team faces, participants were more likely to process a message centrally when the message was presented by their own team face than when presented by another team face—they were more likely to accept a strong argument and less likely to accept a weak argument. In sum, these preliminary data indicate that interacting with an agent wearing one's own team face causes that person to more dedicate more energy towards the task at hand.

These two studies [22, 24] have been utilized solely with two-dimensional avatars in non-immersive displays. Current projects are extending this work to three-dimensional avatars in immersive virtual reality simulations that feature not only the texture being morphed between one or more faces but the underlying shape of the three-dimensional model as well. Previous research has demonstrated that three-dimensional models of a person's head and face built with photogrammetric software is sufficient to capture a majority of the visual features of one's physical self, both in terms of how people treat their own virtual selves [26] and in terms of how others treat familiar virtual representations of others [27].

<h2> Acoustic image

While the majority of research and development in virtual environment technology has focused on the stimulating the visual senses, the technology to richly stimulate the auditory senses is not far behind and possibly holds as much promise in its ability to transform social interactions amongst individuals as does its visual counterpart. Just a few years ago the process to render accurate spatialized (three-dimensional) sound required specialized and expensive digital signal processing hardware. Today, all this processing can be done on consumer-class PCs while easily leaving enough system resources left-over for the user's primary applications. In day-to-day living, we all take spatialized sound for granted just as we take binocular vision for granted. Only when you stop and reflect on the acoustical richness of our natural environments do you realize how much information is derived from the sensed locations of objects: without looking you know from where behind you your colleague is calling your name or that your better quickly step to one side and not the other to avoid being hit by a speeding bicyclist. Spatialization is partly what enables the "cocktail party phenomena" to occur—namely the ability to selectively filter out an unwanted conversation from an attended conversation. As such, our ability to synthetically render these cues in correspondence to three-dimensional visual images enables accurate reconstruction of physical spaces.

More interesting, however, are the possibilities arising from purposely altering the correspondence between the visual and acoustic images. By "warping" relational context, one can hand pick targets that are made maximally available along different channels. Research in cognitive psychology shows that human information processing is capacity limited and that these bottlenecks are largely independent for the visual and auditory channels. This means that by decoupling the visual and auditory contexts one could potentially empower a CVE user with the ability maximize her sensory bandwidth and information processing abilities. For instance, in a meeting scenario one might place two different persons centered in one's field of attention, person A centered visually and person B centered acoustically. This way both A and B could be monitored quite carefully for their reactions to a presentation, albeit along different dimensions.

Just as it is possible to spatialize sound in real-time, it is also possible to alter the characteristics of human speech in real-time. Various software and hardware solutions are available on the consumer market today that can be used to alter one's voice in order to disguise one's identity. While it is not typically easy to transform a male voice into a female voice or vice versa, it is easy to alter a voice with a partial pitch and timbre shift that markedly changes the characteristics so that even someone familiar with the individual would unlikely recognize his identity. The implications of this regarding transforming social interaction are considerable. First, this technology enables the use of duplex voice as a communication channel while still maintaining the anonymity that digital representation allows. Already users in the online gaming community are using this technology to alter their digital personas.

But changing voice to disguise is just one possibility; voice can be transformed in a way that captures the acoustic identity just as the photographs can be morphed to do the same. One form of voice cloning is to sample a small amount of another's voice (e.g., 30 seconds or so) and analyze the frequency components to determine the mean tendencies and then use those statistics to modestly alter the pitch and timbre of your own voice using tools available today. In this way, you could partially transform your voice. While we know of no research that has done so, we believe the end result would be similar to the studies we have discussed in the visual domain. Perhaps a closer analogy to visual morphing is a voice cloning technology recently commercialized by AT&T Labs known as “concatenative speech synthesis.” From a sample of 10-40 hours of recorded speech by a particular individual, it is possible to train a text-to-speech engine that captures the nuances of a particular individual’s voice and then synthesize novel speech as if it came from that individual [28]. While the technology is impressive, it certainly still has a “robotic” ring to it but its potential in CVE use is considerable.

As the next section demonstrates, extending TSI into immersive virtual reality simulations in which interactants’ gestures and expressions are tracked bring in a host of new avenues to explore, and allow for extremely powerful demonstrations of strategies that change the way people interact with one another.

<h1> Transformations of Avatar Behavior

One of the most powerful aspects of immersive virtual reality and, in particular naturalistic nonverbal behavior tracking, is one that receives very little attention. In order to render behaviors onto an avatar as they are performed by the human, one must record in fine detail the actual behaviors of the human. Typically, the recordings of these physical movements are instantly discarded after they occur, or perhaps archived, similar to security video footage. However, one of the most powerful mechanisms behind TSI involves analyzing, filtering, enhancing, or blocking this behavior tracking data in real-time during the interaction. In the current section, we review some previous research in which interactants have transformed their own nonverbal behavior as it occurs, and discuss some of the vast number of future directions for work within this paradigm.

<h2> Non-zero sum gaze

One example of these TSI ‘nonverbal superpowers’ is *non-zero-sum gaze (NSZG)*: providing direct mutual gaze at more than a single interactant at once. Previous research has demonstrated that eye gaze is an extremely important cue: directing gaze at someone (compared to looking away from him or her) causes presenters to be more persuasive [29] and more effective as teachers [30-32]; it increases physiological arousal in terms of heartbeat [33], and generally acts as a signal for interest [34]. In sum, people who use mutual gaze increase their ability to engage a large audience as well as to accomplish a number of conversational goals.

In face-to-face interaction, gaze is zero-sum. In other words, if interactant X looks directly at interactant Y for 80 percent of the time, it is not possible for X to look directly at interactant Z for more than 20 percent of the time. However, interaction among avatars using TSI is not bound by this constraint. In a CVE, the virtual environment is individually rendered for each interactant locally at extremely high frame-rates. Consequently, with digital avatars, an interactant can have his avatar rendered differently for each other interactant, and appear to maintain mutual gaze with both Y and Z for a majority of the conversation, as Figure 5 demonstrates.

NZSG allows a conversationalist to maintain the illusion that he or she is looking at an entire roomful of interactants. Previous research has implemented avatars that use "non veridical" algorithms to drive eye-movements. For example, [35] implemented eye animations that were inferred from the verbal flow of the interaction. In other words, while head movements of interactants were tracked veridically, animation of the eyes themselves were driven not by the people's actual movements, but instead based on an algorithm based on speaking turns. These authors found that the conversation functioned quite well given this decoupling of rendered eye movements from actual eye movements, outperforming a number of other experimental conditions including an audio-only interaction.

Moreover, there has been research directly examining the phenomenon of NZSG. Two studies [36, 37] have utilized a paradigm in which a single presenter read a passage to two listeners inside an immersive CVE. All three interactants were of the same gender, wore stereoscopic, head-mounted displays, and had their head movements and mouth movements tracked and rendered. 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. Moreover, the presenter was always blind to experimental condition; in the augmented condition an algorithm automatically scaled down the magnitude of the presenter's head orientation movements (pitch, yaw, and roll) by a factor of 20 and redirected it at the eyes of both listeners.

Results across those two studies demonstrated three important findings: 1) participants never detected that the augmented gaze was not in fact backed by real gaze, despite being stared at for 100 percent of the time, 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 presenter implementing augmented gaze than a presenter implementing normal gaze.

The potential to use this tool should be extremely tempting across a number of conversational contexts ranging from distance education to sales pitch meetings to online dating chatrooms. Given the preliminary evidence described above, it is clear that avatar-gaze powered by algorithms, as opposed actual human behavior, can be at the very least innocuous, and most likely quite effective, during conversation.

<h2> Digital Chameleons

Chartrand and Bargh [38] describe and provide empirical evidence for the Chameleon effect: when a person mimics our nonverbal behavior, that person has a greater chance of influencing us.

> Such a Chameleon effect may manifest itself different ways. One may notice using the idiosyncratic verbal expressions or speech inflexions of a friend. Or one may notice crossing one's arms while talking to someone else who has his or her arm's crossed. Common to all such cases is that one typically does not notice doing these things—if at all—until after the fact. (p.893)

Data from Chartrand and Bargh's studies demonstrate that when people copy our gestures we like them better, interact more smoothly with them, and are more likely to provide them favor.

Given that typical rendering methods require capturing extremely detailed data concerning their gestures and actions, CVEs lend themselves towards utilizing mimic algorithms at very little added cost. Either from a “nonverbal profile” built from user historical archive data, or from slight adjustments to real-time gestures, it is quite easy for interactants to morph (or even fully replace) their own nonverbal behaviors with those of their conversational partners. There are many motives for interactants to implement the digital chameleon in CVEs, ranging from subtle attempts to achieve influence to powering their avatar with some type of “autopilot” while the user temporarily abdicates his or her seat in the CVE.

Previous research [37] demonstrated that participants often do not detect their own head movements when those movements are rendered at a delay onto other interactants in a CVE. Consequently, to test the digital chameleon hypothesis, Bailenson and Yee [24] ran an experiment in which undergraduate students sat in an immersive virtual environment, at a virtual table, across from an embodied agent. The agent proceeded to read a persuasive passage approximately four minutes long to the participants, whose head orientation movements were tracked while the scene was rendered to them stereoscopically through a head-mounted display. For participants in the *mimic condition*, the agent’s head movements were the exact same movements (on pitch, yaw, and roll) as the participants with a lag of four seconds. In other words, however the participant moved his or her head, the agent mimicked that movement four seconds later. For a separate group of participants in the *recorded condition*, the agent’s head movements were simply a playback of one of the other participants from the mimic condition.

Results of this study demonstrated a huge difference between groups. Agents that mimicked the participants were far more successful at persuading the participants and were seen as more likable than recorded agents. This effect occurred despite the fact that hardly any of the participants detected their own gestures in the behavior of the agents when given a variety of post-experiment questionnaires. These findings are extremely powerful. In order to render the behaviors of an avatar effectively, one must record in high detail all of the actions of the interactants. However, by doing so, the door is opened for other interactants (as well as embodied agents) to employ many types of nonverbal chameleon strategies. In this way, all interactants, some with less than altruistic motives, may achieve a new level of advantage in interaction.

Mimicry is also possible in the auditory channel. Recently, a team at ATR Media Information Science Laboratories in Japan succeeded in doing so [39]. Their idea was to avoid the obstacles of speech recognition and semantics and instead to mimic the overall rhythm and intonation of a speaker. The see if this idea would work, participants were asked to work with an animated agent whom they were told in advance would possess the speech skills of a 1 year-old child. The participants' task was to make toy animals out of building blocks on the computer screen and to teach the agent the names of the toys being built. The agent child would then produce humming like sounds that responded in ways that mimicked the participants' speech rhythms, intonations, and loudness. In a formal study, the levels of mimicry were varied and the effect on the participants' subjective ratings of the agent were then assessed. Ratings were taken that measured cooperation, learning ability, task-achievement, comfort, friendliness, and sympathy. The avatar that mimicked 80% of the time scored highest in user ratings. Just as with the studies reported above on head motions, these findings show that by isolating low-bandwidth dimensions of an interaction it is possible to create a sense of mimicry that does not require a top-down understanding of the interaction.

<h2> Other Behavioral Transformations

There are countless other ways to envision using TSI with the behavior of an avatar. For example, during interaction in CVEs, the automatic maintenance of a “poker face” is possible; any emotion or gesture that one believes to be particular telling can just be filtered out, assuming one can track and categorize that gesture. Similarly, troubling habitual behaviors such as nervous tics or inappropriate giggles can be wholly eliminated from the behaviors of one’s avatars. On the other hand, behaviors that are often hard to generate in certain situations, such as a “genuine smile”, can be easily rendered on one’s avatar with the push of a button.

<h1> Implications and Outlook

The Orwellian themes behind this communication paradigm and research program are quite apparent. Even the preliminary findings discussed in this chapter concerning identity capture, face-morphing, augmented gaze, and digital mimicry are cause for concern, given the huge potential for misuse of TSI by advertisers, politicians, and anyone else whose may seek to influence people via computer-mediated communication. On a more basic level, not being able to trust the very pillars of the communication process—what a person looks like and how they behave—presents interactants in a difficult position. One may ask whether or not it is ethical to keep the behaviors and appearance of your avatar close enough to veridicality in order to prove your identity to other interactants, but to then pick and choose strategic venues to decouple what is virtual from what is real. Is TSI fundamentally different from nose jobs, teeth-whitening, self-help books and white lies?

The answer is unclear. Currently, digital audio streams are “sanitized” over cell phone lines such that the digital information is transformed to present an optimal voice stream using simple algorithms. While this is an extremely mild form of TSI, it is important to point out that very few users of cell phones mind or even notice this transformation. Moreover, the potential ethical concerns of TSI largely vanish if one assumes that all interactants in a CVE are aware of the potential for everyone to rampantly use these transformations.

On a more practical note, an important question to consider is whether or not interactants will bother to pay attention to each other’s behavior if there is no reason to suspect those behaviors are genuine. These strategic transformations utilized in CVEs may become so rampant that the original intent of a CVE—fostering multiple communication channels between physically remote individuals—is rendered completely obsolete. People may completely ignore the nonverbal cues of avatars, given that there is no reason to suspect the cue is genuine. On the other hand, as certain cues become non-diagnostic (e.g., it becomes impossible to infer one’s mental state from one’s facial expression), one can make the argument that interactants will always find the subtle conversational cues that are in fact indicative of actual behavior, appearance or mental state. For example, anecdotal evidence suggests that interactants speaking on the telephone (who do not have any visual cues available) are much more sensitive to slight pauses in the conversation than face-to-face interactants.

CVE programmers may be able to create an extremely persuasive illusion using an avatar empowered with TSI, but will it be possible to mask all truth from an interaction? If there is a lesson to be learned by various forms of mediated communication, it is that people adapt quite well to new technologies. Kendon [40], describes a concept known as *interactional synchrony*, the complex dance that occurs between 1) the multiple channels (i.e., verbal and nonverbal) of a single person during an interaction, and 2) those multiple channels as two interactants respond to one another. Kendon's studies indicated that there are extremely rigid and predictable patterns that occur among these channels during interaction. However, despite this consistent complexity of behavior during conversation, humans are quite adept at maintaining an effective interaction if a channel is removed, for example speaking on the telephone.

Taking away a channel of communication is one thing, but scrambling and transforming the natural correlation among multiple channels is another level of disruption entirely. Transformed social interaction does exactly that, decoupling the normal pairing of behaviors during interaction and, at the whim of interactants, changing the rules of the conversational dance completely. One would expect conversations to completely break down given such an extreme disruption to the traditional order of conversational pragmatics. However, given the results from the empirical investigations of TSI to date, which admittedly are quite limited and preliminary, this has not been the case. Interactants do not seem particularly disturbed by any of the TSI strategies discussed in this paper, and for the most part remain completely unaware of the breakdown among conversational channels.

As future research proceeds, and researchers and systems developers tamper more and more with the structure of interaction, we will provide a true test of the endurance of this conversational structure. One can imagine an equilibrium point in which sufficient amounts of conversational synchrony is preserved, but each interactant is utilizing TSI to the fullest advantage. As systems employing avatars that use these algorithms become widespread, it is essential that this balance point between truth and transformation is achieved. Otherwise, if actions by conversational partners are ships passing in the night, the demise of CVEs and computer-mediated interactions is inevitable.

References

1. Turkle, S. (1995). Life on the screen: Identity in the age of the Internet. New York: Simon & Schuster.
2. Schroeder, R. (2002). Social Interaction in Virtual Environments: Key Issues, Common Themes, and a Framework for Research, in Schroeder R. (ed), The Social Life of Avatars: Presence and Interaction in Shared Virtual Environments, (London: Springer), pp.1-18.
3. Blascovich, J., Loomis, J., Beall, A., Swinth, K., Hoyt, C., & Bailenson, J. (2002). Immersive virtual environment technology: Not just another research tool for social psychology. Psychological Inquiry, 13, 103-124.
4. Slater, A. Sadagic, M. Usoh, R., & Schroeder, R. (2000) Small Group Behaviour in a Virtual and Real Environment: A Comparative Study. PRESENCE: Teleoperators and Virtual Environments 9(1), 37-51.
5. Normand, V., Babski, C., Benford, S., Bullock, A., Carion, S., Chrysanthou, Y., et al. (1999). The COVEN Project: Exploring Applicative, Technical and Usage Dimensions of Collaborative Virtual Environment, PRESENCE: Teleoperators and Virtual Environments, 8(2).
6. Yee, N., chapter in this volume
7. Bailenson, J.N., Beall, A.C., Loomis, J., Blascovich, J., & Turk, M. (2004, in press). Transformed Social Interaction: Decoupling Representation from Behavior and

Form in Collaborative Virtual Environments. PRESENCE: Teleoperators and Virtual Environments, 13 (4).

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

9. Taylor, S.E. & Fiske, S.T. (1975) Point of view and perception so causality, Journal of Personality and Social Psychology, 32, 439-445

10. Benford, S., Bowers, J., Fahlen, L., Greenhalgh, C., & Snowdon, D. (1995). User Embodiment in Collaborative Virtual Environments. In Proceedings of CHI'95 , New York, pp. 242–249. ACM Press.

11. Loomis, J.M., Blascovich, J., & Beall, A.C. (1999). Immersive virtual environments as a basic research tool in psychology. Behavior Research Methods, Instruments, and Computers, 31(4), 557-564.

12. Biocca, F. (1997). The cyborg's dilemna: Progressive embodiment in virtual environments. Journal of Computer-Mediated Communication online., 3.

13. Rheingold, Howard. The Virtual Community: Homesteading on the Electronic Frontier. Revised Edition. Cambridge: MIT Press, 2000.

14. Gibson, W. (1999). All Tomorrow's Parties, Ace Books.

15. Shanteau, J., & Nagy, G. (1979). Probability of acceptance in dating choice. Journal of Personality and Social Psychology, 37, 522-533.

16. Byrne, D. (1971). The Attraction Paradigm. New York: Academic Press.

17. Brock, T.C. (1965). Communicator-recipient similarity and decision change. Journal of Personality and Social Psychology, 1, 650-654.

18. Gaertner, S.L., & J.F. Dovidio. 1977. The subtlety of white racism, arousal and helping behavior. Journal of Personality and Social Psychology, 35: 691-707.

19. Bargh, J. A., Chen, M., & Burrows, L. (1996). Automaticity of social behavior: Direct effects of trait construct and stereotype priming on action. Journal of Personality and Social Psychology, 71, 230-244

20. Blanz, V. & T. Vetter (1999). A Morphable Model for the Synthesis of 3D Faces. SIGGRAPH'99 Conference Proceedings, 187-194.

21. Busey, T. A. (1988). Physical and psychological representations of faces: Evidence from morphing. Psychological Science, 9, 476-483.

22. Bailenson, J. N.,Garland, P., Iyengar, S., & Yee, N (2004). The effects of morphing similarity onto the faces of political candidates. Political Psychology. Manuscript under review.

23. Zajonc, R. B. (1971). “Brainwash: familiarity breeds comfort.” Psychology Today. 3(9): 60-64.

24. Bailenson, J. N. & Yee, N. (2005, in press). Digital Chameleons: Automatic assimilation of nonverbal gestures in immersive virtual environments. *Psychological Science*.

25. Petty, R.E. & Cacioppo, J.T. (1986). The Elaboration Likelihood Model of persuasion. New York: Academic Press.

26. Bailenson, J.N., Beall, A.C., Blascovich, J, Raimundo, M., Weisbuch, M. (2001). Intelligents agents who wear your face: User's reactions to the virtual self . A. de Antonio, R. Aylett, D. Ballin (Eds) Intelligent Virtual Agents 2001, 86-99.

27. Bailenson, J.N., Beall, A.C., Blascovich, J., & Rex, C. (2004, in press). Examining Virtual Busts: Are Photogrammetrically-Generated Head Models Effective for Person Identification? PRESENCE: Teleoperators and Virtual Environments, 13 (4).

28. Lisa Guernsey (July 31, 2001) Software Is Called Capable of Copying Any Human Voice: New York Times: Section A , Page 1 , Column 1.

29. Morton, G. (1980). Effect of eye contact and distance on the verbal reinforcement of attitude. Journal of Social Psychology, 111, 73-78.

30. Sherwood, J. V. (1987). Facilitative effects of gaze upon learning. Perceptual and Motor Skills, 64, 1275-1278.

31. Otteson, J. P., & Otteson, C. R. (1979). Effect of teacher's gaze on children's story recall. Perceptual and Motor Skills, 50, 35-42.

32. Fry, R. & Smith, G. F. (1975). The effects of feedback and eye contact on performance of a digit-encoding task. Journal of Social Psychology, 96, 145-146.

33. Wellens, A. R (1987). Heart-rate changes in response to shifts in interpersonal gaze from liked and disliked others. Perceptual & Motor Skills, 64, 595-598.

34. Argyle, M. (1988). Bodily communication (2nd ed.). London, England UK: Methuen.

35. Garau, M., Slater,M., Bee, S. and Sasse, M.A. (2001). The impact of eye gaze on communication using humanoid avatars. Proceedings of the SIG-CHI conference on Human factors in computing systems, March 31 - April 5, 2001, Seattle, WA USA, 309-316.

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

37. Bailenson, J.N., Beall, A.C., Blascovich, J., Loomis, J., & Turk, M. (2004). Non-Zero-Sum Gaze and Persuasion. Paper presented in the Top Papers in Communication and Technology session at the 54th Annual Conference of the International Communication Association, New Orleans, LA.

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

39. Suzuki, N., Takeuchi, Y., Ishii, K., and Okada, M.: (2003). "Effects of Echoic Mimicry Using Hummed Sounds on Human-Computer Interaction", Speech Communication, Vol. 40, No. 4, pp. 559 - 573 (Jun. 2003).

40. Kendon, A. (1977). Studies in the Behavior of Social Interaction. Indiana University: Bloomington.

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

Figure 1: Non-digital transformations of self utilized currently

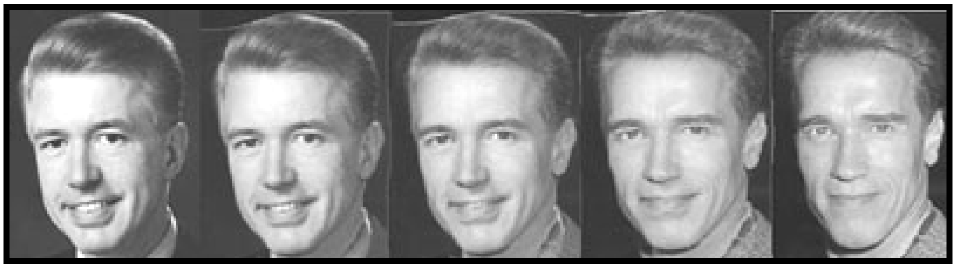

Figure 2: A digital morph of the two-dimensional avatars of Gray Davis (left) to Arnold Schwarzenegger (right).

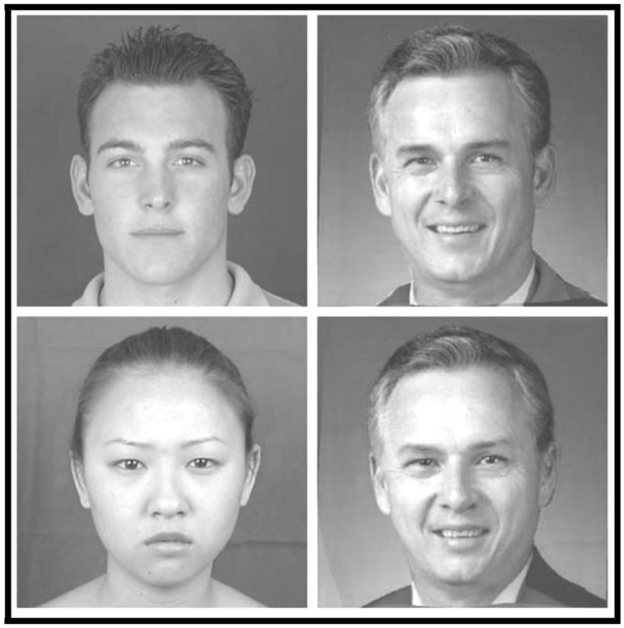

Figure 3: Pictures of the participants are on the left; the blend of 60 percent of an unfamiliar politician and 40 percent of the given participant is on the right.

Figure 4: Four participants (left four panels) and their team face (far right), a morph that includes 25 percent of each of them.

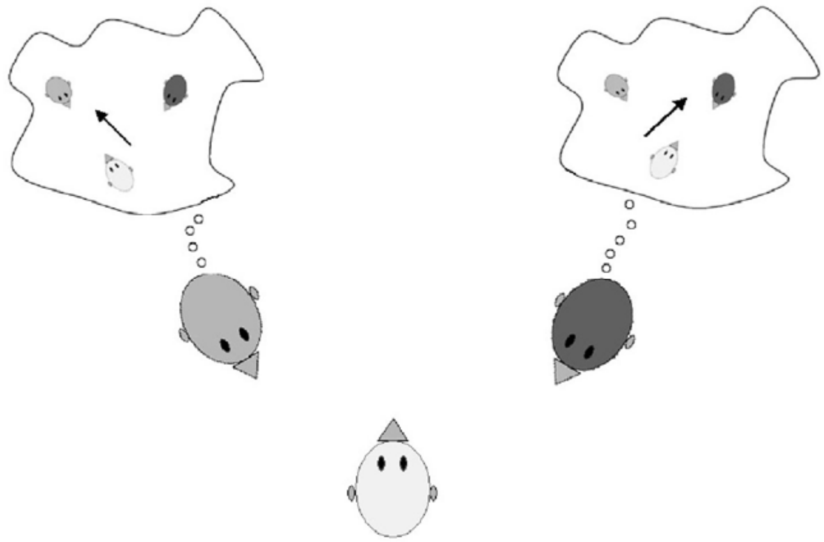

Figure 5: Non-zero-sum Gaze: Both the interactant on the top left and on the top right perceive the sole mutual gaze of the interactant on the bottom.