The Tyranny of Embodiment

(Article for Artifact special issue on Productive Play)

Nick Yee, Palo Alto Research Center

Jason Ellis, IBM T.J. Watson Research Center

Nicolas Ducheneaut, Palo Alto Research Center

During the countercultural movement, the fascination with using technology to trigger sensations of shared disembodied consciousness drove the appeal of LSD, strobe lights, Day-Glo paint, as well as the then nascent virtual reality systems (Wolfe, 1973). As Barlow wrote after his first experience in virtual reality, "Suddenly I don't have a body anymore … the closest analog to virtual reality in my experience is psychedelic" (Barlow, n.d.). Thus, one important aspect of early virtual reality systems was that it allowed us to escape our bodies. This historical perspective on what was novel about virtual environments is revealing because everywhere we look in contemporary virtual worlds we see instead an insistence on creating virtual bodies that mirror our physical ones.

Indeed, some contemporary virtual worlds go to extremes to enable a close-to-perfect reproduction of the human body in digital form. For example, Second Life's factsheet states that "using over 150 unique sliders, [you can] change everything from your foot size to your eye color to the cut of your shirt" (Linden Lab, 2008). The irony is therefore that instead of providing an escape from our bodies, virtual worlds have tended to encourage a meticulous scrutiny and obsessive fascination with them. And along with embodiment comes a wide range of artifacts and practices that cater to those bodies—chairs, houses, vehicles, and even food. In fact, the assumption of embodiment in virtual environments is so pervasive that we oftentimes do not notice a jarring aspect of worlds like Second Life. In a world where people can choose to be anyone they want to be, why do so many Second Life users live in suburban houses and go shopping for Abercrombie and Fitch knockoffs? There are other idiosyncrasies of our contemporary virtual worlds that may also not be obvious due to this assumption of embodiment. For example, if we

took a step back, we might wonder why we need virtual chairs in the first place if our virtual bodies never get tired from standing up.

The metaphor of embodiment and the expectations it brings directs our attention to replicating physical reality in virtual worlds, including the need to create a legion of artifacts that revolve around our bodies. Thus, a virtual meeting is conceptualized as a collection of virtual chairs, virtual tables, and virtual people. But does this insistence on replicating physical reality in virtual worlds limit the kind of work that can be done in virtual worlds? Or put another way, could new forms of work and collaboration be imagined in a world without traditional human embodiment?

There is, of course, nothing wrong with embodiment. After all, one good reason for relying on embodiment is that it provides a host of familiar and well-understood cues for social interaction—mutual gaze, proxemics, gestures, etc. And familiar artifacts, such as chairs, interface with each other to provide a framework of social meaning. Thus, the positioning of chairs around a table facilitates a certain kind of social interaction with certain understood norms of behavior. On the other hand, providing unsituated virtual objects would create an unfamiliar experience requiring overhead to develop social norms and meaning around these novel objects.

Nevertheless, the emergence of art, literacy, and science all hinged on finding alternative modes of representation. For example, music allows us to represent emotions, memories, and experiences in a novel way. Or for example, neither writing nor music employs the metaphor of the human body, but they enable us to think, create, share, and interact with each other in new ways. Moreover, even if we did accept the premise that familiar metaphors easily provide structure and meaning, there are still many other

metaphors that we are familiar with apart from human bodies. Would a brainstorming meeting be more naturally structured by using a representation revolving around plants with its familiar notions of offshoots, branches, maturity, incubation, and cross-pollination?

In the remainder of this paper, we will consider various examples of breaking the traditional insistence on embodiment. These examples include techniques that replace the human body with something else altogether as well as techniques that subtly change the assumed one-to-one correspondence of participants' shared realities. Our goal is not to suggest that these alternatives are inherently superior to a traditionally embodied space, but these examples show us how different our virtual worlds can be. And instead of focusing our attention on how virtual worlds can more faithfully replicate physical reality, we hope that these examples highlight how we oftentimes forget that we can do the impossible in virtual worlds.

The Expectations of Embodiment

There is a set of common expectations that goes along with the insistence on replicating physical reality and physical embodiment. By playing with and deliberately breaking these assumptions in our examples below, we are trying to show how different virtual worlds could be. While not all virtual worlds adhere to these expectations, we feel that most contemporary virtual worlds assume the following.

1) *The expectation of human embodiment.* Users adopt human (or humanoid) avatars in virtual worlds.
2) *The expectation of matched affordances.* Avatars move about and do things the way that people do things in the physical world.

3) *The expectation of congruence.* Users (via their avatars) have different perspectives of the virtual world, but these perspectives are perfectly congruent.
4) *The expectation of single avatar control.* Each user can only control one avatar at a time.

Non-Human Embodiment

The first class of examples revolves around employing non-human forms of embodiment. One rationale for doing so, mentioned above, is that human bodies are not the only metaphors we are familiar with. In other words, other objects may confer novel metaphors for social interaction and work. For example, lurking is a behavior that is possible and quite prevalent (Nonnecke & Preece, 2003) on web-based message forums, but lurking is difficult, if not impossible, in virtual worlds where every user has a visible avatar (barring game worlds where certain classes have the ability to stealth). On the other hand, a virtual tree provides a suitable embodiment for eavesdropping as well as a more appropriate embodiment for people who want to be in a persistent world as an observer. While we might assume that the human body provides the fullest range of interactions, the tree example illustrates that the metaphor of the human body makes certain kinds of interactions impossible. This example breaks both the expectation of human embodiment and matched affordances.

As another example, the canonical virtual meeting typically places each user in his or her own virtual body, but there is no reason why the users couldn't be embodied by the common objects in a meeting instead. The scribe could be embodied by the white board (with access to a suite of tools for note-taking). The discussion leader could be the table and literally change the discussion by changing the shape, size, or arrangement of

chairs around the table. One could imagine a suite of alternate embodiments in a virtual world that might open up a suite of novel metaphors for social interaction and work.

A virtual world populated by alternate embodiments also opens up the possibility for serial embodiment. In other words, the user has no default embodiment and is free to take over, or essentially possess, objects in the world that are not embodying other users currently. Between embodiments, the user would be in ghost form. This notion thus explicitly rejects the status quo of single, static, human embodiment. As Hayles noted, the human body is just an accidental flesh prosthesis (Hayles, 1999). Virtual worlds allow us to confront this accident directly by offering us the possibility of not only alternate embodiment but also serial embodiment.

Borrowing and Amplifying Non-Verbal Cues

A very different approach acknowledges the fact that non-verbal cues are pervasive and important in social interaction. Whether we are dealing with body posture (LaFrance, 1982), eye gaze (Sherwood, 1987; Wellens, 1987), or interpersonal distance (Bailenson, Blascovich, Beall, & Loomis, 2003), non-verbal cues carry a great deal of social meaning. Getting rid of the human body would also mean losing this rich set of cues. Here we discuss several approaches to either borrow or amplify such cues.

One way to use non-verbal cues might be to surreptitiously amplify or hijack them in the background, even as we retain human embodiment. For example, given that eye gaze influences persuasion in social interaction, we might engineer a virtual world where a virtual presenter could maintain eye contact with every member of the audience at the same time (Bailenson, Beall, Blascovich, Loomis, & Turk, 2005). This is possible because every audience member sees the virtual world from their own computer display

and these versions of reality need not be congruent. Thus, this would break the expectation of congruence.

These kinds of strategic filters have been referred to as Transformed Social Interaction (Bailenson, Beall, Loomis, Blascovich, & Turk, 2004). These filters might also be used in the background to enhance social interactions. In the most basic example, the system might insert “please”, “thank you”, or “excuse me” into the chat exchanges between users. Or consider a more complicated example from the social world There where if two users run too close through each other, the system would show each user that the other user walked around them instead (Clanton & Ventrella, 2003). TSI is a strategy that acknowledges that non-verbal cues are too important to discard human embodiment, but at the same time rejects traditional embodiment by leveraging the considerable advantage that the system can have in filtering and manipulating the social interactions between users. In this sense, TSI amplifies interaction possibilities using avatars.

Instead of borrowing or amplifying non-verbal cues, a different tactic might be to overload human avatars with novel cues. For example, in a virtual meeting scenario, participants that haven’t spoken up much may have avatars that become more and more translucent (Bailenson et al., 2008). Or for example, the more negative a user’s words are, the darker his or her shadow becomes. Thus, we can imbue avatars with cues that keep track of important interaction factors that are otherwise invisible but which computer systems can track and display automatically. In this sense, even though human embodiment remains, we can leverage virtual worlds to introduce novel features into a familiar social interaction.

Lastly, another possible approach to distancing ourselves from human embodiment while retaining social cues is to apply familiar non-verbal cues to "non-human" avatars. For example, Babble (Erickson et al., 1999) introduces the notion of a "social proxy." Such proxies provide abstract representations of social interaction that foreground people, proximity, and history. Figure 1 shows a specific social proxy designed to support conversation. The large circle represents a discussion and the colored dots represent people (the people outside the circle are in a different discussion). When an individual speaks, their dot moves to the center of the large circle. As they don't say anything, their dot slowly moves to the edge of the circle. By meaningfully remapping activity to proximity, this social proxy provides an easy way to understand how active a discussion is and *who* is generating that activity. This kind of distillation of social cues from standard online interactions is called social translucence (Erickson & Kellogg, 2000). Moreover, this example foregrounds the fact that a virtual environment need not be in 3D to begin with.

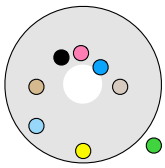

Figure 1: A social proxy

Because social proxies represent people, they beg the question: what is human embodiment? How abstract can a representation be before we consider it non-human? Are the dots in Babble human embodiment? Are Nintendo's Miis? Are emoticons? Indeed, there is a spectrum of representation from highly realistic to more abstract (see figure 2).

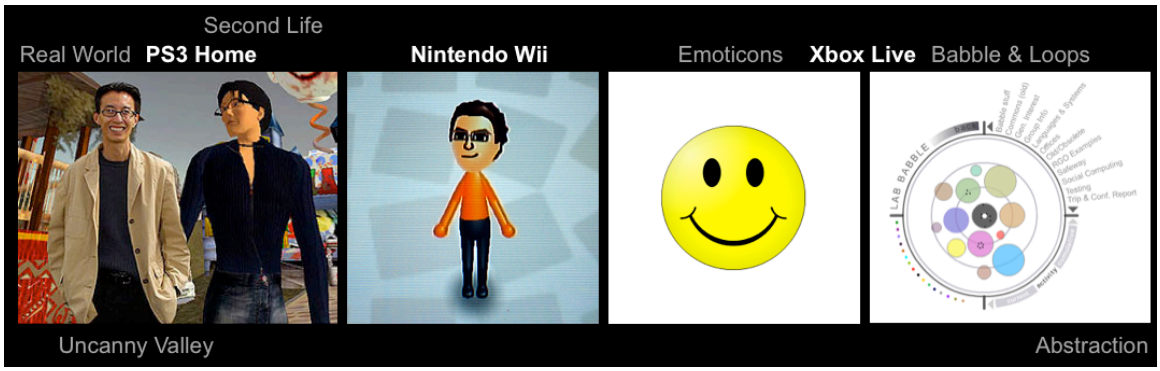

Figure 2: Spectrum of representation

While resolving the question of where human representation ends and non-human representation begins is beyond the scope of this article, perhaps it is safe to say that the notion of human embodiment may be more strongly bound to our perception of each representation than the level of detail in the representation itself. And our perception is deeply informed by the context each representation appears in.

Bodies and Controllers

The typical virtual world avatar parallels a feature of the physical world—every brain only has one body. Virtual worlds however allow us to break this expectation of single avatar control in interesting ways. First of all, one user can control multiple avatars, each with their own viewpoints and performing different actions concurrently. Many video games in the real-time strategy genre employ this version of avatars (e.g., Age of Empires, StarCraft), but this is seldom seen in contemporary virtual worlds. After all, there is no reason why people shouldn't be able to be in multiple places at the same time in a world that allows them to do so. And yet, the one-avatar/one-user expectation is so strong that virtual world users who deliberately challenge it are often frowned upon. In the multiplayer game World of Warcraft, for instance, the practice of "multi-boxing" (that is, using several machines in parallel to control multiple avatars) is considered by

some to be a form of cheating, and it still generates vigorous debates within the players' community.

The reverse is also worth examining. Multiple users can control the same avatar at the same time. Such a configuration might be useful in situations where multiple users can direct their attention towards and control particular aspects of the avatar's behavior to better achieve the goals of the social interaction (Bailenson et al., 2004). For example, one user might focus on the verbal interaction, while another user might focus on emotes and other non-verbal cues, while a third user might be in charge of the private messages. In fact, this kind of virtual presentation might be more powerful than the situation where a single person must handle all of these interactions. The challenge, of course, is assuring the communication done by each individual reinforces that of the others. In some ways, then, this mirrors the challenges of operating complex puppets like the character Jabba the Hutt in Star Wars: Return of the Jedi. Creating the illusion of Jabba required that four puppeteers work in unison: one operating the tail; another controlling the left arm, head and tongue; a third working the right arm, mouth, and reading dialogue; and the final puppeteer controlling the eyes and face via remote control (Wikipedia, 2008). Recent research has also explored how simultaneous user actions in 3D spaces can be harmoniously combined during cooperative object manipulation (Pinho, Bowman, & Freitas, 2002), which suggests that moving beyond a "one avatar-one controller" paradigm can have practical value.

Breaking Rules Productively

At first glance, it may be difficult to imagine how these well-accepted expectations of embodiment can be broken in productive ways, but empirical studies in

some of the areas mentioned have demonstrated measurable benefits. For example, in a study of virtual classrooms where participants (in the role of teachers) were asked to spread their gaze among students, we found that participants performed much better when provided with a supersensory ability (that breaks the expectation of matched affordances)—students literally faded away if they hadn't received gaze from the participant for a while—than in a traditional virtual classroom (Bailenson et al., 2008).

In another study examining how facial similarity could affect voting behavior (Bailenson, Iyengar, Yee, & Collins, in press), we found that participants were more likely to agree with and vote for a political candidate who had taken on 35% of the participants facial features (via photograph morphing). In a virtual presentation setting where each audience member has a different window into the "shared" space, a virtual presenter could adopt 35% of each person's face separately for each audience member. Thus, the expectation of congruence can be broken in powerful ways.

Studies comparing human and non-human embodiments in virtual work settings might also be useful. For example, we may assume that human avatars are natural embodiments for virtual work, but human avatars also invite the need to dress them, to sculpt their faces, to position them correctly, to observe them, and to scrutinize them. In short, human avatars may very well distract people from the actual work itself.

In addition, the direct replication of human bodies can have the side effect of replicating existing social hierarchies. For example, if Bob is the boss in the real world, avatar Bob is still the boss in the virtual world. Studies have shown how status effects can have a negative impact on the productivity of brainstorming sessions (Valacich, Dennis, & Nunamaker, 1992). On the other hand, studies have also shown that manipulating

anonymity in brainstorming sessions improves productivity by eliciting ideas from less vocal group members (Davis, Zaner, Farnham, Marcjan, & McCarthy, 2002). In exploring the affordances of different virtual embodiment schemes, we may very well find that certain configurations improve idea generation while other configurations improve decision-making. In fact, a variety of configurations might be employed for different tasks throughout the process. This ties in with the primary motivation for articulating the assumptions of virtual embodiment in this paper: the insistence on replicating real world representations in the virtual world may prevent us from realizing the full range of work that can be done productively in such worlds.

Ending Thoughts

Our goal in this paper is not to claim that human embodiment is bad, or that breaking the expectations of embodiment is always good, but rather to note that an insistence on human embodiment may distract us from creating new ways of interacting, working, and being. In other words, by forcing virtual worlds to look like physical worlds in form and function, we may be missing out on what virtual worlds may be good for.

Over the past few years, many companies, such as IBM (2007) and Seriosity, have been intrigued by the intersection of online gaming and corporate work. After all, corporate work is becoming increasing virtualized (Ellis, Luther, Bessiere, & Kellogg, 2008) and online gaming oftentimes resembles real work (Yee, 2006). One common misconception that non-gaming executives cling to when observing online games for the first time is that the magic ingredient lies in the 3D—particularly the avatars. Thus, the thinking goes, if only we could put work in a 3D world, it would be more fun. On the other hand, as Farmer noted more than a decade ago however, 3D “isn’t an inherently

better representation for every purpose. 3D is an attribute, like the color blue. Any time you read or hear about how great 3D is and how it's going to change everything about computers and devices, substitute the world blue for 3D" (Farmer, 1996).

Moreover, 3D and avatars function in games in a way that may not be obvious to non-gamers. Avatars are a mechanism that slows down progression in games—a way to provide challenges that keep players from instantaneously completing game goals. This is particularly true in online role-playing games that require a significant time commitment. In such games, players often have to walk a significant distance from point A to point B to complete a quest. If players could complete tasks instantaneously, there would be no game. This is why online gamers can't run through dungeon walls, why they have to accumulate virtual gold for several weeks by killing hundreds of monsters before they can buy a horse that lets them move only 60% faster. Meaningfully slowing down game activities through embodiment can make the events feel more significant for players.

In this light, observing online games to provide guidance on implementing virtual work environments may be dangerous. Well-designed game activities often gain value by extending in-world interactions, while a key metric in judging business operations is the opposite: the faster tasks can be completed, the better. The irony, then, is that one benefit of embodiment in game worlds might well impede users in non-game situations.

Metaphors of embodiment are powerful things. They carry with them an implicit set of expectations. Sometimes, like in online games, the function and consequences of human embodiment and matched affordances (i.e., walking, not being able to walk through walls) may not be obvious at first. More importantly, they may or may not align with the goals of the particular application (e.g., an efficient virtual work setting, a

collaborative classroom). As we explore and develop virtual worlds for a wide variety of applications, it is important to ask whether our insistence on replicating physical reality inadvertently means carrying along unnecessary baggage from the physical world. Indeed, we suggest that it is more fruitful to ask instead what worlds we could create if we broke those expectations purposefully.

References

Bailenson, J., Beall, A., Blascovich, J., Loomis, J., & Turk, M. (2005). Transformed Social Interaction, Augmented Gaze, and Social Influence in Immersive Virtual Environments. *Human Communication Research, 31*, 511-537.

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, 13*(4), 428-441.

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

Bailenson, J., Iyengar, S., Yee, N., & Collins, N. (in press). Facial Similarity between Voters and Candidates Causes Influence. *Public Opinion Quarterly*.

Bailenson, J., Yee, N., Blascovich, J., Beall, A., Lundblad, N., & Jin, M. (2008). The use of immersive virtual reality in the learning sciences: Digital transformations of teachers, students, and social context. *The Journal of the Learning Sciences, 17*, 102-141.

Barlow, J. (n.d.). Being in Nothingness. from http://www.eff.org/Misc/Publications/John_Perry_Barlow/HTML/being_in_nothingness.html

Clanton, C., & Ventrella, J. (2003). *Avatar-Centric Communication in There.* Paper presented at the Stanford HCI Seminars.

Davis, J., Zaner, M., Farnham, S., Marcjan, C., & McCarthy, B. (2002). Wireless Brainstorming: Overcoming Status Effects in Small Group Decisions. *Proceedings of the 36th Hawaii International Conference on System Sciences, 1*, 46-53.

Ellis, J. B., Luther, K., Bessiere, K., & Kellogg, W. A. (2008). *Games for Virtual Team Building*. Paper presented at the ACM DIS 2008 Conference on Designing Interactive Systems.

Erickson, T., & Kellogg, W. A. (2000). Social Translucence: An Approach to Designing Systems that Support Social Processes. *Transactions on Computer-Human Interaction, 7*(1), 59-83.

Erickson, T., Smith, D. N., Kellogg, W. A., Laff, M., Richards, J. T., & Bradner, E. (1999). Socially Translucent Systems: Social Proxies, Persistent Conversation, and the Design of "Babble". In *Proceedings of CHI 99*.

Farmer, R. (1996). Now in 3-D! *Wired* Retrieved October 31st, 2008, from http://www.wired.com/wired/archive/4.01/farmer.if.html

Hayles, K. (1999). *How We Became Posthuman: Virtual Bodies in Cybernetics, Literature, and Informatics*. Chicago: University of Chicago Press.

IBM. (2007). Virtual Worlds, Read Leaders: Online games put the future of business leadership on display. Retrieved October 31st, 2008, from http://www.ibm.com/ibm/ideasfromibm/us/giogaming/073007/index.shtml

LaFrance, M. (1982). Posture Mirroring and Rapport. In M. Davis (Ed.), *Interaction Rhythms: Periodicity in Communicative Behavior* (pp. 279-298).

Linden Lab. (2008). Factsheet: The Technology Behind the Second Life Platform. Retrieved October 31st, 2008, from http://lindenlab.com/pressroom/general/factsheets/technology

Nonnecke, B., & Preece, J. (2003). Silent participants: Getting to know lurkers better. In D. Fisher & C. Lueg (Eds.), *From Usenet to Cowebs: Interacting with social information spaces*: Springer Verlag.

Pinho, M., Bowman, D., & Freitas, C. (2002). Cooperative Object Manipulation in Immersive Virtual Environments: Frameworks and Techniques. *Proceedings of VRST 02, 1*, 171-178.

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

Valacich, J., Dennis, A., & Nunamaker, J. (1992). Group size and anonymity effects on computer-mediated idea generation. *Small Group Research, 23*, 49-73.

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

Wikipedia. (2008). Jabba the Hutt: Concept and Creation. Retrieved October 15, 2008, from http://en.wikipedia.org/wiki/Jabba_the_Hutt#Episode_VI:_Return_of_the_Jedi

Wolfe, T. (1973). *The Electric Kool-Aid Acid Test*. New York: Farrar Strauss and Giroux.

Yee, N. (2006). The Labor of Fun: How Video Games Blur the Boundaries of Work and Play. *Games and Culture, 1*(68-71).