

You can be too rich: Mediated communication in a virtual world

Greg Wadley, Martin R. Gibbs

The University of Melbourne
Parkville, 3010, Australia

{greg.wadley,martin.gibbs}@unimelb.edu.au

Nicolas Ducheneaut

Palo Alto Research Center
Palo Alto, CA, USA

nicolas@parc.com

ABSTRACT

Internet-based virtual worlds (VWs) have emerged as a popular form of collaborative virtual environment. Most have offered only text chat for user communication; however several VWs have recently introduced voice. While research has demonstrated benefits of voice, its introduction into the popular VW Second Life (SL) was controversial, and some users have rejected it. In order to understand the benefits and problems that voice brings to virtual worlds, we used qualitative methods to gather data from SL users and analyse it. We discuss our results in the light of media-richness theory and its critiques, arguing that preferences for voice or text reflect a broader problem of managing social presence in virtual contexts.

Author Keywords

virtual worlds, Second Life, voice, media richness

ACM Classification Keywords

H5.1. Information interfaces and presentation (e.g., HCI):
Multimedia Information Systems.

INTRODUCTION

Collaborative virtual environments (CVEs) use multiple media to enable rich interaction over a distance. Their defining feature is the simulation of a 3-dimensional space, in which users are embodied as avatars which can be positioned like bodies in physical space. Three-dimensionality affords users the semantics of place, including deixis, indexical language, and body orientation. By allowing users to simulate aspects of offline interaction such as how they position themselves when conversing (Yee et al., 2007), these systems aim to enable mutual awareness and a natural articulation of collaboration (Benford et al., 1994).

While avatars allow a limited body-language, linguistic communication requires a mechanism for transmitting text or audio between users. Following the 3d metaphor, the rules of message transmission can be based on avatar location. For example, messages might be sent only to users whose avatars are near the sender's avatar, simulating the transmission of sound in air and enabling "a natural intuition about mutual audibility" (Smith et al., 2002). Alternatively a CVE can allow conversation

between users at arbitrary virtual locations, simulating the use of telecommunication devices within the virtual space (Wadley et al., 2005).

"Desktop" virtual worlds (VWs) delivered via consumer-grade PCs and the Internet have emerged as a popular genre of CVE for work and recreation. VWs initially offered only text-based communication, but in recent years several have included facilities for transmitting speech. A leading VW, *Second Life* (Ondrejka, 2004), introduced voice in 2007, with mixed success. A recent vendor press release (May 2009) stated that while voice has been welcomed by some, half of SL's users had not adopted it.

In the study reported here we interviewed users, analyzed forum discussions and conducted participant research, in order to understand the benefits and problems of voice in Second Life. We discuss our results in the light of media-richness theory and its critiques and research into contexts of technology use.

Media richness

The concept of media richness pervades discourse about virtual worlds. This theory (Short et al, 1976; Daft and Lengel, 1986) proposed that the richer a medium - the more it transmits the 'social presence' of collaborators - the more effectively it should substitute for face-to-face interaction. A rich medium conveys not just language symbols, but also information about the interlocutors themselves. Some VW users have expressed a desire for increased richness: for example Saeed et al. (2008) discovered that educators' intention to use Second Life was predicated on it. Schroeder (2002) noted that in the context of virtual environments, media richness is often called 'co-presence': the degree to which users feel they are together in the same space.

Some studies have highlighted the low fidelity of the simulation provided by VWs. For example diminished visual cues create problems for conversational turn-taking (Nilsson et al., 2002) and reference to objects (Hindmarsh et al., 2001). Researchers are developing new input-output mechanisms that tie users' virtual-world representations more closely to their physical-world actions. Salinas (2002) studied the use of devices that permit communication by touch. Logitech have a camera which translates users' facial expressions onto their avatars. SL's vendor Linden Lab has demonstrated a means for controlling avatar actions via users' body movements (<http://www.handsfree3d.com>).

Undoubtedly the technical properties of a medium

influence its ability to convey information about interlocuters. However the notion of a simple relationship between media-richness and either choice of medium or communication style has been much critiqued. Data contradicting the expectation that people should choose the richest medium available has led researchers to propose alternative models. The social-influence (Fulk et al., 1990) and critical-mass (Markus, 1990) frameworks point to the influence of collaborators. Communication richness depends on users' prior experience with the medium and with each other, and their intentions (Carlson and Zmud, 1999; Walther, 1992). Researchers have focused on the benefits of using a collection of media (Watson-Manheim and Belanger, 2007). It is likely that personality type influences preference for rich or lean media (Goby, 2006). People wishing to engage in "impression management" (Carlson et al., 2004), or who are shy (Stritzke et al., 2004), may choose text for its low social presence. People value mediated communication for the ability it offers to manage self-presentation (O'Sullivan, 2000), especially in face-threatening situations (Feaster, 2008). Control over self-presentation is not only desired in recreational systems but is actively exploited in work contexts (Vast, 2007).

Therefore while all communication media are in a sense replacements for face-to-face communication, the usefulness of a medium cannot be reduced to the fidelity with which it emulates face-to-face (Hollan and Stornetta, 1992). This applies equally to VW use. Churchill and Bly (1998) studied use of a MUD within a workplace, concluding that: "*visually oriented media richness is not a prerequisite for the creation of sufficient social co-presence for maintaining collaborative relationships.*" Moore et al. (2006) recommended that avatars in game worlds should reflect their owners' use of the UI but not necessarily their physical context, noting that "*Ultimately designers must consider the trade-offs between privacy and transparency and decide what degree of each is most appropriate given the kind of gaming experience they wish to create.*"

Contexts of communication

HCI researchers have focused increasingly on the context in which use happens. This is in part a reaction to the movement of computing from the workplace desktop into social and domestic spaces (Howard et al., 2007). Online and offline activities and identities intertwine as communicators multitask within their physical and virtual contexts (Aarsand, 2008). IRC users engage in multiple simultaneous conversations, paying careful attention to who can read which utterances. Text chat offers "different sets of 'mutual monitoring possibilities' ... different ways in which they allow us to be present to one another and to be aware of other peoples' presence" (Jones, 2002). They also control the visibility of conversations to physically co-located others by tilting screens or hiding communication windows behind other applications. In a text conversation "the 'muting' of the aural mode allows them to carry on on-line conversations which are inaccessible to others who are physically co-present" (Jones, 2002).

Online game-players using voice want to know who can overhear their transmissions (Gibbs et al., 2004). Telecommunication users attempt to control their projection of personal information. Eavesdropping was a common concern in the days of party-line telephony (Fischer, 1994). Mobile phone users choose SMS to avoid eavesdropping, and use features of their phones to control social interactions (Madell and Muncer, 2007).

Voice in VWs

An apt illustration of the variance in desire to project social presence is the existence of users who exploit the communicative leanness of existing VWs in order to role-play. Researchers have differentiated *immersionists* who use VWs for identity exploration from *augmentationists* who project their offline identity into the VW (Boellstorff, 2008). Many recreational VWs are designed so that users are visually represented as a character in a fictional universe: lean media allow role-play to extend beyond avatar appearance and into user communication.

For this reason the implementation of voice in VWs has prompted debate. Voice projects information such as a user's gender, ethnicity and age, interfering with the anonymity which many value (Raybourn, 2002) and which is exploited during role-play. Thus, it has been argued, voice communication breaks one of the fundamental purposes for being in the VW (Bartle, 2003). While research has demonstrated benefits of voice in game worlds (Williams et al., 2007), many SL users seem to have rejected it. This is a concern for vendors of these large systems.

METHODS

To study attitudes and practices related to voice communication in SL we interviewed users, convened in-world discussions, analyzed forums and blogs, and conducted participant research into both recreational and workplace use of SL.

We interviewed 10 experienced users from a variety of backgrounds who used SL for activities including socializing, teaching, business and art. Interviews were open-ended and lasted one to two hours. Some took place in SL, using voice or text according to the interviewee's preference, while some were face-to-face.

We moderated a focus-group about voice within a long-running weekly SL discussion group. This lasted an hour and comprised about twenty participants, including long-term SL users and commentators. We monitored the official SL user forum, two popular email-lists, and several prominent blogs during the research period, saving relevant threads for later analysis.

The researchers were active users of SL and other VWs throughout the research period and were involved in several working groups which held regular meetings in SL. These included a geographically-dispersed research group within a large institution, and a conference organizing committee. This required that we deal with problems of voice and text communication in our own work practice. We also organized a weekly in-world research seminar series that utilized both voice and text.

One author took part in several mixed-reality music performances. This period of intensive participant research allowed us to develop insights and cross-check experiences reported by informants.

Using grounded methods we coded data into themes which reflect the concerns of our participants. These are presented in the next section.

RESULTS

Voice vs text: range of preferences

The topic of voice elicits extreme opinions and there is a sense that individual SL users either love voice or hate it. Users offered opinions from "I hate voice so badly. I just hate it. I want it to go away" to "It hugely enhances SL for a lot of people". Others felt that extreme pro- or anti-voice positions were inappropriate and that individuals should be left to choose the medium they preferred. One interviewee felt that voice worked in some situations, with text always available as a fallback, and was happy to let others make the choice.

There was speculation about how many people were using voice in SL. Many felt that it was restricted to isolated pockets or to certain types of users. Some forum posters said they rarely encountered voice-enabled avatars. Discussion participants agreed: "Predominantly, everyone uses text". However a Linden Lab employee in the discussion said that: "People do like voice. Some people really don't, but a lot of people really do."

Richness praised

Participants reported that speech conveyed more richness and subtlety of meaning, more quickly, than did typed text, and were often positive about this. One interviewee praised: "the immediacy of the connection, and the intensity of the bonds/emotional commitments", and reported that "voice does help with understanding some of the communication overtones". A forum poster perceived text as too "formal" and unable to convey some kinds of meaning. Others felt that voice was more suited to conveying complex concepts when helping fellow users: "Especially when explaining something to a newbie it is faster and easier to speak. Meetings with friends and group members are more efficient when most people are talking instead of typing. "

However one interviewee felt that voice "broadens, but reduces the depth of the communication", offering the analogy that "this is the difference between a movie and a book." He felt that in the development of relationships "the addition of voice can be both powerful and restricting."

One informant felt that: "SL's original silent world is more magical without the crassness of voice and interruptions by real humans"

Identity and anonymity

Preference for anonymity is one of the main reasons cited by SL users for rejecting voice. An interviewee commented: "You don't feel so exposed in text". Another said: "for me SL is all about creating the image that

people want to be, and to be honest I would rather imagine how some people talk than actually hear them."

Playing an avatar of opposite gender is common in SL, but one's offline gender is usually exposed in speech. "I also have many SL friends who role-play, and don't want to reveal identity with their voice. ... For example I have several friends who are male but choose female avs. They obviously don't use voice." One male interviewee with a female avatar said he had been apprehensive about the introduction of voice to SL to the point of becoming anxious and depressed.

For female users, voice can cause problems by giving away the fact that they really are female: "What turned me against it was the fact as soon as guys would hear a girl's voice it was like a feeding frenzy." On the other hand, one forum poster found that in some situations she was treated better if she used voice. One male interviewee who conducted content-creation work as a female avatar reported being treated well by a client who assumed he was a young woman.

One forum poster pointed out that actors and other offline role-players successfully assume fictional identities while using their natural voice: "I was in the Society for Creative Anachronism for years and was quite able to maintain immersion in the medieval world in First Life." (This illustrates that identity-play and acting are not the same thing.)

Voice-refusers sometimes encountered suspicion from other users who assumed they were trying to conceal their identity. "You get the oh it's really a man, woman, kid. People assume all the wrong things when someone doesn't use it not when they do." One text-user reported being expelled from a piece of land by its voice-using owner who assumed the visitor must be concealing their identity. Another suggested: "I think there needs to be an accepted community standard way of saying 'I don't use voice' which is accepted as a choice and not thought of as suspicious." . A forum poster who attempted to join a group reported that they demanded they all first voice-chat, to ascertain whether he was trustworthy. Some of our participants preferred such transparency of identity. An interviewee who used SL for conversing with business partners and customers said she liked voice as it gave her a better idea of who she was talking to.

Some preferred to use voice only when they already knew the people to whom they were speaking. A forum poster reported that they typically only used voice with family and friends. One interviewee "would tend to approach new people with text first in social settings".

Several informants referred to the 'immersionist' and 'augmentationist' divide. An interviewee pointed out: "voice makes SL a better telephone to chat with real-life friends. But what use is voice to a furry?" (Members of the 'furry' sub-culture present as animals instead of people.) Another referred to "two different SL audiences - recreational and business, with very different uses and needs. Business prefer / need voice." (Some studies have shown that this is not necessarily the case, e.g. El-Shinnawy and Markus, 1998). Some forum posters

perceived a connection between the vendor's introduction of voice and its management of these two user groups. "I suspect there's always been tension between the people who want to extend their RL selves into the virtual space, making it like 'Skype-3D', and those who want to take something other than RL there. [...] In the early stages of Second Life it was the second group that were hooked and were consequently paying; but in order to finance the next stage of its growth Linden Lab need to provide features which interest the first group at the risk of upsetting the second."

Personality

Several informants felt that a user's personality contributed to their preference for voice or text: "I find that a lot of people who don't like voice are shy and lack confidence." Some felt that voice allowed offline extroverts to dominate online (cf. France et al., 2001): "If you introduce sound the whole ambience changes: the shy are revealed as shy, and the noisy start to dominate. (It is hard to type LOUDER than everyone else in the room!)". An educator said: "Text is sometimes much more appealing for those who are a bit more introvert, whilst voice can enable the performers amongst us." She felt that "sometimes text is a more effective mode of communication and participants will share more." An interviewee felt that: "People are insecure about their voices - they don't like the recorded sound of their own voice, worse than seeing their own photo."

Specialized use-cases suited to voice

Our informants described some specific scenarios to which voice is particularly suited. Several were involved in education, and all of these felt that voice was suited to online teaching. One said: "I use both voice and text. Has been fabulous having the option to use voice chat added to SL. Particularly for education. Many educators I work with didn't 'get hooked' until voice was added."

One interviewee used SL for teaching a foreign language and liked being able to both type a word and speak it. In this scenario he found no problems with voice other than technical ones. A discussion participant agreed: "There is a language school in SL. They *have* to use voice."

Live music is popular in SL and the voice channel is an easy way to transmit it. Forum posters described encounters with users playing guitar, singing, and performing theatre, poetry, storytelling and comedy. A discussion participant noted: "a choir playing live in SL. That's hard to do in text!" Another commented: "I just came from a play rehearsal that couldn't have happened without voice."

SL is unusual among VEs in that it allows users to create content. Some forum posters felt that voice was well suited to collaborative building: "I am with a group of builders working on a quite big project. I usually use voice for our meetings and other communications." A blog poster commented: "I like to use voice when I am hanging around and talking to friends while building, because it frees up my hands to continue creating things."

One interviewee attended regular interest-group meetings

using voice, and another held business meetings. A discussion participant who used SL to telecommute said that: "Voice is a lifesaver. It makes meetings much better." A blog poster commented on the use of voice for cybersex: "As for who is using voice, it's escorts!"

The processability of text

SL users value the processability of text. For example, a text conversation can be recorded, copied to another application or searched for keywords. An interviewee commented: "It's easy to keep minutes of a meeting if it's in text". A forum poster agreed: "Everyone has full transcripts of exactly what was said and by whom, and what was agreed to. No misunderstandings. In voice there's no way to record or transcribe the conversation, or to identify who is speaking". A discussion participant felt that "text will persist just because it is easy to remember what was said". Another noted: "I attended a lecture where the speaker used voice. And, in that context, it was really good to HEAR her rather than read her. Shame I have no transcript though".

Some reported exploiting the opportunity to edit text messages in the short gap between composing and sending: "Sometimes I go to type something but then delete it, because I realise I should not say it."

Multi-tasking

Informants felt that text was better suited to multitasking and conducting simultaneous conversations. A forum poster noted: "When your main mode of conversation is text, you can talk to lots of people at once, about different things, and the participants don't all know you're talking to other people. When it's voice you all talk about the same thing, in public, usually with the same people." An interviewee said: "especially when you're multitasking as I often am, text is easier." A discussion participant noted: "When you're on a phone call, all your attention is focused on the call. That's normal: voice demands high bandwidth. Whereas on text, you can multitask pretty easily." An educator praised multitasking: "I WANT my students to be texting, IMing and generally doing bunches of other stuff".

Text in SL is presented as a real-time unthreaded series of messages, similar to a chat client. An interviewee said: "With text, I can go back to the log and follow various threads of chat". A forum poster reported that: "I find that typed meetings go faster, because you don't have to wait for people to finish talking, you don't have to repeat what you just said, you don't need 'rules of order'." Another interviewee reported being confused about who was talking when he heard several people using voice at once. Another felt that: "It can be difficult to have a meeting or any interaction with a lot of people present as one person's voice tends to cut through someone else's and people don't naturally take turns when talking." Likewise, "voice in games suffers from the inability of players to see each other and use visual cues to decide who will speak next." These observations support the findings of Herring (1999).

It is common in SL discussion groups, seminars and

similar forums for a main speaker to use voice while other attendees type comments and questions. An interviewee said: "I have lead discussions where I have deliberately used both modes [...] Perhaps presenters use voice to share at the beginning of session then ask people to work in groups with text." A forum poster reported that: "in both SL and in academic Skype conferences I have recently found myself in private side-chats, discussing how and why the main thread was going wrong, planning an intervention with some allies, and then intervening" (see also Yankelovich et al., 2005). On the other hand, one interviewee found voice to be confusing when speakers were also IM-ing each other.

Group size

Echoing Lober et al. (2007), informants reported that the larger the group, the better suited is text. One interviewee said the largest group he had used voice with was six users, and that he'd felt like he was talking in a crowded room. A discussion participant said: "Voice works well in small intimate groups and is very good at one-to-one's. Its efficacy diminishes as the group gets bigger".

Usability

Most people speak fluently, but few are competent typists. A forum poster said: "I use voice whenever possible. I usually don't mind whether or not the other person types or talks as long as I can avoid having to type :)". An interviewee said that "typing skills makes a huge diff". Another commented on the overloading of manual controls: "My favorite use is for keeping hands free while exploring the world with a friend, since it's the only way to walk and talk simultaneously." Some commented on accessibility problems: "One of my family members in game has cerebral palsy which makes typing difficult for her." (Of course, some disabled users cannot use voice.)

However users complain of frequent problems getting voice as it is currently implemented in SL to work properly. Problems stem from multiple sources: voice is not standard on PCs and requires extra hardware to be purchased and kept in working order, it is analogue in nature and must be converted to digital for transmission, admitting the possibility of noise, and there are many points in transmission where loudness can unintentionally vary (eg microphone placement).

Problems with proximity-based voice

Most SL users operate voice in its proximity mode, which simulates the transmission of sound in air by attenuating volume with distance. One interviewee felt that while people have a good understanding of how far their voice travels in the physical world, the rules of transmission in SL are less intuitive. Another felt that SL's implementation is not realistic enough, in that one can't control transmission distance by shouting or whispering.

By mimicking the physics of real-world voice transmission, proximity voice reintroduces real-world problems. One forum poster noted that if someone attends a public lecture or large meeting in SL, it is hard to be in a location where one can hear all the speakers at once. They might position themselves to hear the main speaker

clearly, but would then be unable to hear an audience member at the back ask a question. SL has no 'megaphone' to project voice over a large area.

Unintended transmission

A side-effect of using microphones is that sounds other than the user's voice may be sent. A forum poster complained: "I don't do voice, and keep it turned off. [...] The last couple of people I listened to on voice not only were hacking their lungs out at me, but making some other disgusting sounds as well." Another reported: "Most the time it's some mom screaming at the kids with the mic on, someone having their music on very loud in their RL house". An interviewee who created audiovisual art in SL was concerned that voice recreates a problem found in physical galleries and music venues: audience members chatting and disturbing the listening experience of fellow attendees.

Some informants reported maintaining awareness of which sounds they might transmit, and adapting their practice to suit. A forum poster said: "I don't do voice. Too high an ambient noise level in the room, with my kid and/or my mate watching TV or playing music in the same room." Another commented: "I have family at home and I'm constantly being called by name to do something. That's one major thing I don't want broadcast over the internet."

Similarly, some users do not want their virtual-world conversations to disturb those physically co-present: "I usually do not use voice, mainly because it is distracting to others in the house." A forum poster noted: "Hell, maybe you have a baby sleeping next to you. Or a housemate who doesn't want to listen to you talk into a mic while she's watching a movie." A parent remarked: "I'm not much of a TV fan, but I've found an acceptable "alone together" compromise where they watch and I play, and we occasionally interact. Headphones and microphone don't really mix in that environment."

Some are wary of eavesdropping: "They may have family members in close proximity or be in circumstances where talking to a computer isn't private enough." SL affords eavesdropping in the virtual world by users whose avatars are near the speakers' avatar. This has no direct analogy in media which do not simulate space. "My partner and I regularly talk to each other in-world using voice. And although I have an eye on the mini-map and an awareness that people can listen in I still do it." SL users can also listen from their movable, invisible view-camera: "People can zoom in from a long way away to listen in. For newbies all of that is very hard to understand. "

Griefing and etiquette

Some people try to make the Internet experience unpleasant for others. Voice seems to create new ways of doing this. One interviewee claimed to have switched SL voice on when it was implemented, left it on for ten minutes, then turned it off after hearing only verbal abuse. She felt grievers were worse in voice than in text. Another claimed that voice is primarily used for griefing: "The only places I really see it in use are at the Welcome

Areas -- and usually to harass and belittle others or say racist/sexist stuff". A discussion participant who plays a 'furry' agreed: "Every time I hear someone do voice in SL I just hear people insult me and say Furfag and the like".

Informants offered different reasons for why grieving was worse when spoken rather than typed. One felt that an insult "becomes more personal when voice is used", implying a 'dark side' to media richness. Some thought voice attracted people who were disposed to grieving. Some felt voice exaggerated offline problems: "The girls stop talking completely, the shy people shut up mostly, and all that is left are the 12-18 year old guys, and it becomes a locker room". It is possible that griefers prefer voice simply because it makes them less likely to be caught. Voice is harder than text to log and it is unlikely that Linden Lab systematically monitor it.

Voice lends itself to a particular form of grieving reported by one informant: "The wildest game I've seen played with voice chat is when someone records 'choice phrases' from someone, uploads the WAVs, and plays them back while the person is still there. [...] Don't voice drunk, sarcasm is your enemy, and watch what you say around welcome area crowds."

While grieving represents deliberate misuse, our informants recognized that some problems were accidental. Some felt that over time the development of etiquette would improve practice: "I think we're all still working on the social niceties of what's ok and when".

DISCUSSION

You can be too rich

Our results demonstrate that Second Life users - like other communicators - rarely want to project maximum social presence into the virtual world. These users negotiate complex combinations of physical and virtual contexts, each with the potential for eavesdropping and unintended transmission. Using 'lean' media for linguistic communication within the VW helps them cope with this. Therefore media-richness is a contestable design goal for virtual worlds, as it has proven to be for other forms of CMC.

Much of the appeal of recreational virtual environments stems from the opportunity they afford users to play characters in a fictional world. Even in environments without a back-story, such as Second Life, people frequently create representations of themselves that are idealizations of their real selves (Ducheneaut et al., 2009), and are concerned to control the amount and type of presence they convey. The immersionist / augmentationist divide can be understood as a spectrum of different preferences for media-richness. Immersionists want to be 'in' the virtual world, and in it with other people, but do not want to 'be themselves' in it: rather they want a 'fictional presence'. Forcing users to project personal information is at odds with this desire.

The augmentationist / immersionist categorization suggests that virtual worlds should either emulate reality with maximum fidelity, or be fenced off by a 'magic circle' and understood as separate realities with their own

culture. But these two extremes may not fit much actual use. We argue that most people adopt a third approach which is something of a hybrid of these two: they exploit the social opaqueness of communication media as a resource in their interactions with others (Erickson and Kellog, 2000; Gibbs et al, 2004). They understand which parts of their bodily and situational reality are on display and which are not, and use this knowledge to construct a performance (in the sense of Goffman) to achieve social goals. Communicators are aware that their use of 'socially lean' media is less vulnerable to being monitored. They exploit this to multi-task and 'multi-context'; engaging in multiple conversations and impression management with collaborators both mediated and physically co-located.

Similarly it has been shown that SL users exploit the detachable view-camera to 'park' their avatars in socially-acceptable orientations while moving their focus of attention elsewhere (Irani et al., 2008; Wadley and Ducheneaut, 2009). Such actions illustrate a politics of attention (Turner and Reinsch, 2007), and lean media make it possible to give the impression that appropriate attention is being paid.

This Goffman-esque view need not imply deception of a malicious kind. VW users may indeed feel themselves to be co-present and cooperating in a shared environment, while maintaining an awareness of what is and is not being transmitted about their physical context. They use this 'reality gap' to engage in mundane activities such as answering the phone or eating, without interrupting the task at hand. These offline activities need not be accountable for effective online cooperation to take place. It is often assumed that virtual environments should be as immersive as possible, that using a CVE should be like being transported to another place, and that these systems succeed when users are so absorbed in the simulation that they forget their physical surroundings (Biocca et al., 2003). But most communicators need to multi-task: an employee using a VW at work, a parent engaging simultaneously in family life and a Second Life discussion, or a teenager playing games while chatting with friends, cannot afford to be so immersed in the online world that they ignore offline events.

Flexible richness

Our data illustrate the wide range of users and use-cases that virtual world designers must cater for, and the concomitant variance in communicative richness suited to each scenario. How can a VW provide this variance? The only answer to date seems to be to offer both voice and text. But our data show this is inadequate: recall that some users do not wish to participate in SL if *anyone* is using voice, and that conversely a preference for text, when voice is available, may be viewed with suspicion. Perhaps cultural change is needed so that these different groups understand and trust each other's preferences. Ethnographic research can contribute to this.

It is interesting to note that voice chat remains surprisingly untouched by technological innovations that might make it more useful and/or palatable to VW users. For instance, it is reasonably straightforward to modulate

an audio stream's pitch to disguise or enhance a person's voice (the current YouTube craze around Auto-Tune¹ is a good illustration). Bartle (2003) suggested this could be pushed even further in the form of a 'voice mask' that VW users might use to sound younger or older, switch their gender, or generally alter any part of their voice they are uncomfortable with in a public, digital space. This might provide a compromise between lean text and rich voice.

As the example of electronic mail illustrates, and unlike what Media Richness Theory might predict, 'digitizing' communication media transforms their affordances. Email is much more than 'electronic postal mail': this, however, became clear only as users experimented with the medium, re-conceptualized its use, and as the underlying technology (email clients) adapted accordingly. Due to its very recent addition to VWs, it is quite possible that voice chat remains in the initial phase of a similar adoption curve. As such, there is an opportunity for more research and design on voice technology that goes "beyond being there" (Hollan and Stornetta, 1992).

CONCLUSION

Our results emphasize that virtual worlds are a unique context for communication. Complexity arises out of the ability of users to conduct multiple simultaneous conversations in both physical and virtual contexts, each with its own physics of visibility and audibility. The user experience can become a maelstrom of impression management, identity play, and uncertainty over what is being transmitted to whom.

Voice profoundly impacts people's experience of virtual worlds. The controversy over its introduction reflects a broader tension over how much information about the people behind avatars should be projected into virtual space. Maximum social presence is rarely desired: rather users want to tailor the presence they transmit according to circumstances. This is no different to how people behave in physical environments (Goffman, 1959).

Researchers have already emphasized the need to design for flexible use. It seems one challenge now for virtual worlds is how to design for flexible richness.

ACKNOWLEDGMENTS

We would like to thank the participants in this study for giving up their valuable time, and the various colleagues who have discussed this work with us.

REFERENCES

Aarsand, P.A. Frame switches and identity performances: Alternating between online and offline. *Text & Talk* 28, 2 (2008), 147-165.

Bartle, R. *Designing Virtual Worlds*. New Riders: Indianapolis (2003).

Benford, S., Bowers, J., Fahlén, L.E., Mariani, J.A. and Rodden, T. Supporting Cooperative Work in Virtual Environments. *The Computer Journal* 37, 8 (1994),

653-668.

Biocca, F., Harms, C., and Burgoon, J.K. Toward a more robust theory and measure of social presence: Review and suggested criteria. *Presence* 12, 5 (2003), 456-480.

Boellstorff, T. *Coming of Age in Second Life*. Princeton University Press, Princeton (2008).

Carlson, J.R., George, J.F., Burgoon, J.K., Adkins, M. and White, C.H. Deception in computer-mediated communication. *Group Decision and Negotiation* 13, 1 (2004), 5-28.

Carlson, J.R. and Zmud, R.W. Channel expansion theory and the experiential nature of media richness perceptions. *Academy of Management Journal* 42, 2 (1999), 153-170.

Churchill, E. F. and Bly, S. It's all in the words: Supporting work activities with lightweight tools. *Proc. of SIGGROUP* (1998), 40-49.

Daft, R.L. and Lengel, R.H. Organizational information requirements, media richness and structural design. *Management Science* 32, 5 (1986), 554-571.

Ducheneaut, N., Wen, M-H, Yee, N. and Wadley, G. Body and Mind: A Study of Avatar Personalization in Three Virtual Worlds. *Proc. of CHI* (2009).

El-Shinnawy, M. and Markus, M.L. Acceptance of communication media in organizations: richness or features?. *IEEE Transactions on Professional Communication* 41, 4 (1998), 242-253.

Erickson, T. and Kellogg, W.A. Social translucence: An approach to designing systems that support social processes. *ACM Transactions on Computer-Human Interaction* 7, 1 (2000), 59-83.

Feaster, J. *Interpersonal Media Use in Face-Threatening Situations and the Construct of Information Control*, Annual Meeting of the International Communication Association (2008).

Fischer, C. S. *America Calling: A Social History of the Telephone to 1940*, University of California Press: Berkeley (1994).

France, E.F., Anderson, A.H. and Gardner, M. The impact of status and audio conferencing technology on business meetings. *International Journal of Human-Computer Studies* 54, 6 (2001), 857-876.

Fulk, J., Schmitz, J. and Steinfield C. W. A social influence model of technology use. In Fulk, J.S. and Steinfield, C. W. (eds.) *Organizations and Communication Technology*, Sage Publications, Newbury Park (1990), 117-140.

Gibbs M., Hew, K. and Wadley, G. Social Translucence of the Xbox Live Voice Channel. *Proc. of International Conference on Entertainment Computing* (2004).

Goby, V. P. Personality and Online/Offline Choices: MBTI Profiles and Favored Communication Modes in a Singapore Study. *Cyberpsychology & Behavior* 9, 1 (2006).

Goffman, E. *The Presentation of Self in Everyday Life*. Anchor: New York (1959).

¹ <http://en.wikipedia.org/wiki/Auto-Tune>

- Herring, S. Interactional coherence in CMC. *Computer-Mediated Communication* 4, 4 (1999).
- Hindmarsh, J., Fraser, M., Heath, C. and Benford, S. Virtually missing the point: configuring CVEs for object-focused interaction'. In Churchill et al. (eds.): *Collaborative Virtual Environments: Digital places and spaces for interaction*, Springer, London (2001), 115-139.
- Hollan, J. and Stornetta, S. Beyond Being There. *Proceedings of CHI'92* (1992), 119-125.
- Howard, S., Kjeldskov, J. and Skov, M.B. Pervasive computing in the domestic space. *Personal and Ubiquitous Computing* 11 (2007), 329-333.
- Irani, L. C., Hayes, G. R. and Dourish, P. Situated practices of looking: visual practice in an online world. *Proc. of CSCW* (2008).
- Jones, R.H. The Problem of Context in Computer Mediated Communication. *Georgetown Roundtable on Language and Linguistics* (2002).
- Lober, A., Schwabe, G. and Grimm, S. Audio vs. chat: The effects of group size on media choice. *Proc. of HICSS* (2007).
- Madell, D.E. and Muncer, S.J. Control over Social Interactions: An Important Reason for Young People's Use of the Internet and Mobile Phones for Communication?. *Cyberpsychology & Behavior* 10, 1 (2007).
- Markus, M.L. Toward a "critical mass" theory of interactive media. In Fulk, J. and Steinfield, C.W. (eds). *Organizations and Communication Technology*, Sage Publications: Newbury Park (1990), 194-218.
- Moore, R., Ducheneaut, N. and Nickell, E. Doing Virtually Nothing: Awareness and Accountability in Massively Multiuser Online Worlds. *Computer Supported Cooperative Work* 16, 3 (2006), 265-305.
- Nilsson, A., Heldel, I., Axelsson, A-S., and Schroeder, R. The Long-Term Uses of Shared Virtual Environments: An Exploratory Study. In R. Schroeder (ed.) *The Social Life of Avatars*. Springer: London (2002), 112-126.
- Ondrejka, C. R. Escaping the Gilded Cage: User Created Content and Building the Metaverse. *New York Law School Law Review* 49, 1 (2004), 81-101.
- O'Sullivan, B. What you don't know won't hurt me: Impression management functions of communication channels in relationships. *Human Communication Research* 26, 3 (2000), 403-431.
- Raybourne, E. M. Designing an emergent culture of negotiation in collaborative virtual communities: the DomeCity MOO simulation. In Churchill et al. (eds.): *Collaborative Virtual Environments: Digital places and spaces for interaction*. Springer: London (2002), 247-264.
- Saeed, N., Yang, Y., and Sinnappan, S. Media richness and user acceptance of Second Life. *Proc. of Ascilite* (2008).
- Salinas, E-L. Collaboration in multi-modal virtual worlds: comparing touch, text, voice and video. In Schroeder, R. (ed.) *The Social Life of Avatars*. Springer: London (2002), 172-187.
- Schroeder, R. Copresence and interaction in virtual environments: An overview of the range of issues. *Proceedings of Presence 2002*, 274-295. (2002).
- Short, J., Williams, E. and Christie, B. *The social psychology of telecommunications*. Wiley, London (1976).
- Smith, R. B., Hixon, R., and Horan, B. Supporting flexible roles in a shared space. In Churchill et al. (eds.): *Collaborative Virtual Environments: Digital places and spaces for interaction*. Springer: London (2002), 160-178.
- Stritzke, W.G.K., Nguyen, A. and Durkin, K. Shyness and computer mediated communication: A self-presentational theory perspective. *Media Psychology* 6, 1 (2004).
- Turner, J.W. and Reinsch, N.L. The Business Communicator as Presence Allocator: Multicommunicating, Equivocality, and Status at Work. *Journal of Business Communication* 44, 1 (2007), 36-58.
- Vast, E. The presentation of self in a virtual but work-related world: from protagonists to fools. In K. Crowston and S. Sieber (eds.) *Virtuality and Virtualization*. Springer: Boston (2007).
- Wadley, G. and Ducheneaut, N. The out-of-avatar experience: object-focused collaboration in Second Life. *Proc. of ECSCW* (2009).
- Wadley, G., Gibbs, M. and Benda, P. Towards a framework for designing speech-based player interaction in multiplayer online games. *Proc. of Australasian Conference on Interactive Entertainment* (2005).
- Walther, J.B. Interpersonal effects in computer-mediated interaction: a relational perspective. *Human Communication Research* 19, 1 (1992) 50-88.
- Watson-Manheim, M.B. and Belanger, F. Communication mode repertoires: dealing with the multiplicity of media choices. *MIS Quarterly* 31, 2 (2007), 267-293.
- Williams, D., Caplan, S. and Xiong, L. Can you hear me now? The impact of voice in an online gaming community. *Human Communication Research* 33, 4 (2007), 397-535.
- Yankelovich, N., McGinn, J., Wessler, M., Kaplan, J., Provino, J. and Fox, J. Private communications in public meetings. *Proc. of CHI* (2005), 1873-1876.
- Yee, N., Bailenson, J. N., Urbanek, M., Chang, F., and Merget, D. The unbearable likeness of being digital: the persistence of nonverbal social norms in online virtual environments. *Cyberpsychology and Behavior* 10, 1 (2007), 115-121.