

Usability and Sociability of the Xbox Live Voice Channel

Kevin Hew

Department of Information Systems
The University of Melbourne
Parkville Vic 3010, Australia
+613 8344-1500

kevinhew@hotmail.com

Martin R. Gibbs

Department of Information Systems
The University of Melbourne
Parkville Vic 3010, Australia
+613 8344-1394

martinrg@unimelb.edu.au

Greg Wadley

Department of Information Systems
The University of Melbourne
Parkville Vic 3010, Australia
+613 8344-1586

gwadley@unimelb.edu.au

ABSTRACT

As part of a project exploring the adoption of online videogames, we used a Model of Technology Appropriation to understand users' initial reaction to, and use of, the voice communication channel provided by Xbox Live.

We found that although users expected voice to be an advance over text-based communication, in practice they found voice difficult to use. In particular, users experienced difficulties controlling the voice channel. These difficulties led some users to reject voice as a mode of communication. We also found that the headset created an interesting dynamic between users co-located in the room with the console and those 'present' via the Internet.

There appear to be usability and sociability problems with the way the voice channel is currently configured in some Xbox Live games. We argue that game developers will need to address these problems in order to realize the potential of voice in online multiplayer videogames.

Categories and Subject Descriptors

H.5.3 [Information Interfaces and Presentation]: Group and Organization Interfaces – *collaborative computing, computer supported collaborative work, synchronous interaction.*

General Terms

Design, Human Factors

Keywords

Computer games, Xbox Live, Computer Supported Cooperative Play, Voice communication, Usability, Sociability

1. INTRODUCTION

The video game industry is valued at US\$30 billion a year worldwide [19]. Online video gaming is expected to be worth \$US4.3 billion by 2005 (Forrester Research, quoted in [13]), and is considered by some to represent the future of the computer gaming industry [13]. Launched in the US in 2002 and in Australia in October 2003, Microsoft's *Xbox Live* is a networking system allowing multiplayer gaming via *Xbox* videogame consoles. It appears that Microsoft and other vendors of online console networks intend social interaction to be an important part of the online gaming experience, and have implemented features such as voice communication and identity management in their products to support this.

Xbox Live combines a novel cluster of features in the one product, including: multiplayer gaming over the Internet; console rather than PC platform; a voice communication system with microphone and earphone headset; and a 'gamertag' system for the central management of online identity. In particular, and of interest in this paper, *Xbox Live* combines

voice communication with a virtual environment, connecting users via broadband Internet for the purpose of playing console-based video games.

Previously, online communication has typically relied on typed text messages, and text is still the dominant medium for communication in computer games. However, *Xbox* consoles do not include a keyboard, and users of *Xbox Live* are able to communicate with other players by voice in real time. This new mode of interaction makes *Xbox Live* an interesting technology to study. In this paper we address the following research question:

How do users react to their first encounter with voice communication in *Xbox Live*?

To collect data we used a constructive interaction technique for usability testing [23], and focus group discussions, to gauge users' initial reactions to voice in multiplayer, online console videogames. Our analysis was informed by a social-constructivist understanding of technology adoption and appropriation developed by Carroll *et al.* in their Model of Technology Appropriation [5-7, 12].

In the following section, we outline concepts generated by considering online multiplayer video-gaming as an instance of Computer Supported Cooperative Play (CSCP)*, and discuss their relevance to understanding computer mediated voice communication such as that found in *Xbox Live*. We then present the basic framework of Carroll *et al.*'s Model of Technology Appropriation (MTA) before outlining our research design. Our results are then presented.

We found that while participants expected voice to be an advance over text-based communication, in practice many found voice difficult to use. However, in our testing we also found that the implementation of voice using a 'hands-free' microphone and earphone headset created an interesting communication dynamic between two 'social spaces': one within the room of a user's console, and one within the online game that was being played. We conclude with a discussion of our findings and their implications for the design and configuration of voice communication in online multiplayer video games.

2. Computer Supported Cooperative Play

Xbox Live is a salient example of Computer Supported Cooperative Play (CSCP). CSCP has recently been defined as the 'mutual engagement by two or more individuals in recreational activity mediated by a computing environment' [28]. While some CSCP activities appear to be competitive, these can be viewed as cooperative interaction in the pursuit of

* Thanks to Connor Graham for coining this term in its current usage.

recreation. CSCP builds upon concepts developed in studies of Computer Supported Cooperative Work (CSCW) and uses them to explore the use of computer mediated collaborative technologies in a recreational context, such as multiplayer video games.

As a discipline, CSCW has examined how groups of people work together, and the role that technology can play in helping them collaborate with one another to complete work-related activities as a collective [10]. Important concepts within CSCW include computer mediated communication (CMC), collaboration and coordination, and shared environments. CSCP adds to this set of concepts Ray Oldenburg's 'third place' as a metaphor for places of play. Oldenburg [22] describes the concept thus:

Third places exist on neutral ground and serve to level their guests to a condition of social equality. Within these places, conversation is the primary activity and a major vehicle for the display and appreciation of human personality and individuality. Third places are taken for granted and most have a low profile. The character of a third place is determined most of all by its regular clientele and is marked by a playful mood.

After the home and the work place, third places are where informal public life occurs. They nurture social ties outside of the home and workplace and are characterized by the acceptance of a diversity of patrons engaged in playful and convivial conversation. They serve as a landmark where people can connect with others in the community.

Examples of 'physical' third places include pubs, cafes and bookstores. A virtual environment might be considered a third place if it is a shared public space, in which diversity and equality are promoted, and whose primary purpose is social interaction between the people who use the space. Examples of 'virtual' third places might include MUDs, MOOs, UseNet discussion forums, and multiplayer role-playing games [25, 27].

The 'third place' is a useful metaphor for understanding CSCP and therefore online multiplayer gaming. Two important concepts emerging from the metaphor are explored in this paper: sociability and convivial communication media. Our claim is that these concepts – previously defined and explored across several heterogeneous research domains – assist in answering the research question posed in this paper and allow us to account for the way users reacted to, and used, the voice communication in *Xbox Live* during their first encounters with this technology. These concepts also assist in understanding the phenomenon of CSCP more generally. In the following sections we discuss these concepts further, with a focus on the voice communication medium of *Xbox Live*.

2.1 Sociability

The concept of sociability is understood as being 'concerned with planning and developing social policies and supporting social interactions' [24: p605] in the context of designing for online communities. Sociability and usability are closely related yet usefully separated concepts for analyzing and designing technologies that support online communities. While usability concerns the interaction between a user and a technological artifact, sociability concerns the interactions between people that occur via the artifacts. Sociability includes the policies and social norms as well as the design features of the mediating technology that govern and influence behaviour online.

As Bannon has argued [1], consideration of the social layer of computing systems is crucial in system design. This is further supported by Kutti's critique of HCI [15], identifying the need to understand human-computer interaction in terms of social factors as well as the task-human-artifact dynamic. We believe sociability is an important construct for understanding CSCP, and acknowledge the importance of designing policies and technology to enable cooperative sociable interaction.

Sociability emphasizes the importance of not only planning for the interaction between the player and the game, but also planning for the interactions between people within the game. While sociability concerns the interactions between players in an online game, the mediating technology must be usable in order to support the desired social interactions.

Xbox Live has a number of features directly designed to manage the sociability of the system. These include a fixed user-selected pseudonym ('gamertag') for player identity and accountability, player rankings for many games, and a friends list to allow players to quickly find players with whom they prefer to play. In addition, *Xbox Live* has introduced a voice channel for communication between players, and Microsoft has mandated that all *Xbox Live* games implement voice communication. The configuration of the voice channel – the manner in which it is implemented within a particular game – and its resulting usability as a convivial medium, will significantly impact on the sociability of games within the *Xbox Live* network. Issues associated with CSCP and voice as a computer mediated communication medium are discussed in the next section.

2.2 Communication Medium

In a 'third place', conversation is the primary activity. Turkle [27] describes examples of how, in the past, online text-based games such as MUDs have been co-opted as places of conversation transcending game-oriented goals.

Online videogames could be co-opted in the same way. Unlike most existing online gaming systems, Microsoft's *Xbox Live* system incorporates voice communication at its core. Microsoft requires developers to include voice in their games, and consumers, in order to join the *Xbox Live* network, must purchase a connection kit for their console, which includes a headset with headphone and microphone. The co-availability of voice and 3D graphics in consoles offers the potential of a rich communication medium, as described by Daft and Lengel [9], to support social interaction.

Information richness is the ability to 'overcome different frames of reference or clarify ambiguous issues to change understanding in a timely manner' [9]. Media richness is the capacity of a communication medium to convey rich information [9]. Rich media are able to convey more complex and ambiguous messages. Voice is often regarded as a richer medium than text, due to its capacity to provide immediate feedback between conversation participants, and due to the presence of additional cues such as those found in the intonation and rhythm of speech. It is also worth noting that the ability to conduct rich communication is not solely dependant on the communication medium itself [20]. The context of the message, the shared culture of the participants, and the history of the communication also play a role in the richness of the communication. Thus, the virtual environment in games as well as the shared culture of regular players contributes to players' ability to engage in rich communication.

At first glance, voice communication would appear to be a useful communication medium for a variety of CSCP tasks such

as tactical coordination in team-based action games. Voice communication may come to the fore especially in online role playing games (RPGs), which rely fundamentally on player interaction and conversation. On the other hand, some industry commentators and game designers have recently disputed this assumption and opined that voice conversation will detract from the immersion experience of RPGs precisely because of its media richness [2].

However, in terms of CSCP activities, voice communication can be a useful addition to the game playing experience and the sociability of the gaming environment if it enhances the amount of socially significant information available to users. That is, in order to support CSCP, and the cooperative and collaborative tasks associated with this kind of computer mediated play, a communication medium also needs to be ‘socially translucent’. ‘Social translucence’ [11] refers to a design approach that emphasizes the importance of making socially significant information visible in a CMC environment. Three characteristics of socially translucent systems are visibility, awareness and accountability. By making people and their actions visible to other people in the system, users become aware of others, and accountable for their actions. Awareness of others brings social rules into play which govern people’s actions [11]. Visibility makes people accountable for their actions by the fact that ‘I know that you know that I know’ what you have done. Thus in the design of CSCP systems, it is important for users to be able to easily identify who they are interacting with, and to be able to readily map activity and actions within the game environment to particular game identities and personas.

3. MODEL OF TECHNOLOGY APPROPRIATION

Because we are examining how people respond to the new voice communication system offered by *Xbox Live*, we have framed our questions and analyzed our data within a model of technology appropriation. We chose the model of Carroll *et al.* [5 - 7, 12], a social-constructivist, user-centered model developed to investigate mobile phone use among young people. Carroll’s Model of Technology Appropriation (MTA) describes a continual process of appropriation through time, from adoption or rejection upon initial exposure to ‘technology as designed’, through appropriation or disappropriation after extended use, until final integration of the technology into daily life (‘technology in use’).

Our focus is on *Xbox Live* as a recreational communication technology that may support online communities. Carroll’s model was designed to study a ‘lifestyle’ communication product (mobile phones), rather than a task-focused workplace technology, and should be well-suited to studying the appropriation of *Xbox Live*.

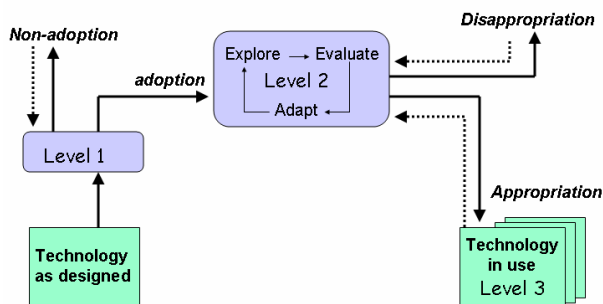


Figure 1. The Model of Technology Appropriation (adapted from [6])

Research into online computer games has primarily focused on experienced users [18, 27, 29]. The MTA suggests that use of a communication technology will change with users’ experience. The MTA describes appropriation as a process, fluid between three levels:

Level 1 represents a user’s initial evaluation of a product, resulting in a decision to adopt or reject, based on the technology’s perceived attractors and repellers.

Level 2 represents a user’s deeper evaluation of the technology through use. This is similar to the concept of a consumer’s post-purchase evaluation of a product [8, 17] during which the consumer evaluates the product through ownership and use. A negative evaluation causes the user to disappropriate the technology.

Level 3 represents a user’s long-term integration of a technology into their daily life. A user may adapt or augment a technology and so transform its original design. A user may also change their practices in response to the appropriated technology.

Use of the MTA requires collection of qualitative data through methods such as observation, questionnaires, focus groups and user journals, to identify attractors and repellers of the technology at each of the three levels. Because *Xbox Live* became available only recently, our research focused on level 1 of the MTA.

4. RESEARCH APPROACH

Our research was conducted with six groups of three volunteers. Each session was conducted with one group at a time, and consisted of a laboratory observation followed by a focus group discussion. The sessions were approximately two hours in duration. At the start of each laboratory observation, participants completed a background questionnaire. They were then asked to play multiplayer games with other players over the Internet using *Xbox Live*. We observed their interaction both with other players via the network, and with other participants in the room. After the session, a researcher led the participants in a focus group discussion. Open-ended questions sought the participants’ opinions about their first encounter with *Xbox Live*. Questions then focused on topics such as the voice headset, and playing with people online. The sessions were recorded using a single video camera directed at the participants. This data was analyzed to identify common themes across the groups through a process of open-coding, then closed-coding to refine and confirm prevalent themes.

4.1 The Participants

Most of the participants were aged between 18 and 24 years. Participants in one group were all over 24 years old. Participants on average reported that they played PC-based computer games more often than console games (6 hours per week versus 2 hours). Two participants reported that they played computer games for over 20 hours per week. Only three participants reported that they played console games more often than computer games.

Participants liked several different types of game. The most popular genres were first-person shooters (13 participants played games in this genre), action games (13), sports (10), strategy games (10) and role-playing games (9).

All participants owned or regularly used a PC and had Internet access. Only three participants had broadband internet access at home. All but one owned a mobile phone.

4.2 Games Used

The two games mainly used during the research were *MotoGP* and *Unreal Championship*.

MotoGP is a motorbike racing game that is packaged with every *Xbox Live* starter kit. *MotoGP* incorporates a ‘lobby’ area, outside the main racing arena, where players waiting for their race to start can talk to each other in a party-line fashion. During a race, *MotoGP* uses a proximity-based algorithm for voice communication, allowing players to talk to each other if they are close together on the track.

Unreal Championship is a first-person shooter allowing either ‘all against all’ or team-based play modes. In all against all play, the voice channel is implemented as a party-line, where all players in the game can hear all others speak. In the team games, all players on a given team can speak to and hear each other.

4.3 Use of Groups in Laboratory Observation

Groups were used in the observation sessions to encourage participants to express their views and describe their experience. This approach was based on the constructive interaction technique developed by O’Malley *et al.* [23] for usability testing. Also known as co-discovery learning [14], this technique aims to create an environment in which participants verbalize their thoughts more readily than does an individual ‘thinking aloud’ to a researcher [21].

Another reason for using small groups was to create a laboratory-based situation analogous to the social setting in which console games are often played. As Jarett *et al.* [13] pointed out, ‘When you ask console gamers about their favorite style of play, most will respond that they like playing on the couch with their friends.’ That is, console gaming is often a social activity involving a small number of physically co-located people, and our research design attempted to reproduce this.

4.4 Focus Groups

A focus group discussion after each game playing session provided participants with an opportunity to discuss and reflect upon their initial encounter with *Xbox Live*. Focus group discussions are appropriate when the goal is to discover people’s opinions of an experience, idea, or event [16]. Focus groups can provide information about perceptions, feelings, and attitudes of users by ‘tapping into the human tendency to develop attitudes and perceptions relating to concepts, products and services, or programs in part by interaction with other people’ [16]. Group discussions also allowed a natural connection between the group game experience and the ensuing discussion

5. FINDINGS

The primary purpose of connecting consoles via the Internet is to allow people to play games with others who are not in the same room. This increases the opportunity to play with other people rather than with computer-generated opponents. The participants in our study all valued the social interactions entailed in playing with other people. According to one participant,

It’s the interaction with other people and it’s the buzz playing with other people and if you’re really competitive and you beat the other person you feel good about it.

The participants in our study cited three major reasons why they preferred to play against people rather than computer-generated opponents: unpredictability, fairness, and enjoyment of competition. Human players were regarded as being unpredictable because they are creative and do not follow a pre-programmed ‘script’. Some of our participants felt that the computer ‘cheated’ or had an unfair advantage over them, particularly at high difficulty settings, because the computer opponents could be unfairly fast, accurate, or omniscient. Participants reported deriving more satisfaction from beating human opponents than computer-generated ones, ‘even though you can’t see the person.’ Knowledge that there was another person in the game, even though they were not co-present, substantially increased enjoyment of the game.

Clearly users want to play against other people. One could assume that any feature that enhanced the social interactions of online multiplayer gaming would increase users’ enjoyment of the system. We would expect the addition of voice communication to gameplay to enhance social interaction. However the voice system, as it was configured in the games our participants played, often seemed to detract from the enjoyment of the game as a social experience. In the following sections we report on users’ reaction to voice in *Xbox Live*.

5.1 Usefulness of Voice Communication – Perception vs. Practice

Participants expected that voice communication would enrich multiplayer gaming. They felt that voice was a step forward from the traditional text-based messaging found in games:

The headset is good in tactical swat team games. That’s really good because that’s really fun. Like at LAN parties and stuff, you shout across computers and stuff like that.

In particular, many of our participants thought that voice would be beneficial in tactical team-based games, by better supporting coordination of team members:

I quite like the headset, the fact that you can talk. Especially in *Unreal* where, well the fact that you’re supposed to work as a team and you can sort of do that quite well. I think it’s better than having to type or anything like that.

An important advantage participants cited for using the voice headset was the freeing up of hands from the task of typing messages. According to one participant: ‘Voice is way better than text because text is annoying, its slow, you have to hit “i” to type it in.’ This would allow players to maintain control over their character’s movements while they talk to other players.

[Voice] was good because with network games on computers you have to type messages and that interferes with the game. But, [voice] was good because it didn’t interfere with what you were doing.

This assumed advantage created by voice over other forms of in-game communication was mentioned by the majority of participants in the study. We, as researchers, certainly expected voice to be a positive augmentation of the gaming experience. However, many participants stopped using the voice headset during observation sessions, effectively rejecting the technology. In practice, it would seem, our participants found voice communication to be unhelpful and not particularly useful or useable in the games used in this study.

Participants experienced a number of difficulties with the use of voice. These difficulties can be understood as being problems of channel control: controlling what is received, and what is

sent, over the voice channel. We discuss these problems in the following two sections.

5.2 Controlling What is Heard Through the Voice Channel

Participants reported two problems with what they heard through the headset: ‘noise’ over the voice channel, and an inability to identify who was talking.

Participants felt that a lot of what was said over the voice channel resembled noise more than conversation. ‘It’s background noise. This is like background music rather than having a real conversation.’ Four types of noise were present on *Xbox Live*.

First, the voice channel included speech that was not intended for the listener. Participants could hear conversations that they were not a part of. Due to the design of the voice communication, these conversations were just as prominent in the voice channel as were conversations directed at the participants.

Second, participants overheard conversations that were apparently between groups of users co-located around a distant console on the network. By leaving the microphone active, part of that conversation would be broadcast over the network. As one participant commented, ‘It just sounds like they are talking to themselves. I mean, we [gestures to other participants in the room] are talking to ourselves.’

Third, participants encountered non-conversational sound that was unintentionally transmitted by other users, such as television, people making motorbike sounds, and background music.

Fourth, sometimes the non-conversational sound encountered by participants appeared to have been sent intentionally, to limit the usability of the channel for others. Some of our participants referred to this as ‘spam’. Examples included loud incoherent speech, relentless trash-talking, noise-making devices such as mobile phones, and music at such a volume that the sender’s headset microphone had apparently been held close to a hi-fi speaker.

In response to noise of this kind, one participant suggested a preference towards typing over voice. He reasoned, ‘that bit of effort in typing text just means you don’t get crap on the airwaves.’ That is, the ease with which noise could be generated over the voice channel was seen as a negative feature once users had experienced a variety of irritating and anti-social examples of its usage.

In addition to the problem with noise, participants had problems identifying who was talking on the voice channel. Participants were not able to link the voice they heard via the headset to either the name of the user or the user’s avatar on the screen:

It’s hard to identify who you are talking to. I mean you can’t get the sense of community if you are connected to all these people, you can’t really see them or you don’t know who you are talking to.

This was particularly pronounced in *Unreal Championship*, where the voice channel was configured in a broadcast, party-line mode, such that everyone on the same team could hear each other equally, regardless of their position in the game world. This effectively ‘disembodied’ the voice, because the volume and clarity of transmission was unrelated to proximity. *MotoGP* uses a coarse proximity-based algorithm that switches voices on or off according to their proximity on the track. As a result voices suddenly appeared, mid-sentence, with no sense of a

person approaching or receding. Participants found these voices just as disembodied as those in *Unreal Championship*.

Clearly the ability of online gamers to connect socially with one another will depend on their ability to identify who is currently talking. However it is also important for a user to know who is listening.

5.3 Controlling What is Sent Over the Voice Channel

Participants experienced two problems involving transmissions they made over the voice channel: controlling who could hear what they said, and knowing whether their utterances were heard.

When participants spoke into the headset microphone, they could neither control nor know who was listening to their transmission. Participants desired the ability to direct their messages to specific people, but could not find this functionality. One said, ‘I didn’t know who I was talking to. There wasn’t functionality to select who I was talking to.’ And another, in a different session, stated: ‘One thing I didn’t like was not knowing who you could or couldn’t speak to.’

Participants experienced the voice channel as chaotic and out of control. One participant suggested:

It might be good if you could direct the speech, so you can specify before you talk who you want to hear it and who you don’t want to hear it.

Suggestions of this kind were made frequently by participants.

A related problem was participants’ uncertainty over whether their utterances were being heard by the intended recipients. Participants often did not receive a response from other players in the game, and were uncertain whether they were being heard. One participant was heard to say: ‘I’m hanging around the flag. Where are you? Can you hear me? Hello?’ During other sessions participants repeatedly asked: ‘Is anyone there?’ or ‘Hello?’ to prompt a response. Participants received no visual indication as to whether their transmissions were being sent or heard successfully.

5.4 In-room Multiplayer Versus Online Multiplayer

Xbox Live is designed to allow users to connect to others over the Internet. A form of multiplayer gaming already available on consoles is playing with friends located in the same room as the console. Participants noted the possibility of combining the two types of multiplayer gaming: ‘If we all could play at the same time, against the other people, that would be really good.’ This would create a connection between two distinct social spaces: one based on physical co-location, the other on Internet communication.

During the observation sessions participants were clearly able to move comfortably between these two social spaces. It was common to see participants speak over the network, then immediately afterwards speak to someone in the room, while continuing to play the game. Situations in which this happened included:

- When participants paused during an *Xbox Live* conversation to express their opinions about the system to the other participants in the room;
- When another participant in the room suggested that they ask someone on the network what to do next;

- When a participant conveyed what they had heard on *Xbox Live* to the other participants in the room.

Sometimes a participant would pause during a local conversation to respond to something that was said on *Xbox Live*.

While participants found it easy to conduct a conversation simultaneously with people in the room and on the network, they typically left their headset and microphone on while doing this, thus broadcasting their 'local' conversation to the network. They appeared not to consider the fact that people on the network could hear what they were saying. While this could create network noise as discussed earlier, it could also allow for the non-intended transmission of private information to other game players on the network.

The *Xbox Live* headset has a mute button to prevent local conversation being broadcast. However it was more common to see participants move the microphone away from their face, or take off the headset, rather than press the mute button. Whether these strategies are effective in stopping broadcast cannot be known by the user.

There was evidence of participants adapting the headset's intended use, to allow for more than one local person to interact with online users. Examples included:

- One participant used the handset while another participant used the headset;
- A participant not wearing the headset would talk louder in order to transmit over the network.

6. DISCUSSION

6.1 Usability Problems Hindered the Social Interaction in Games

The findings suggest that the potential to engage in computer mediated social interaction during the observation sessions was hindered by the problems faced when participants used *Xbox Live*. As mentioned in the findings, many participants rejected the voice headset, due to their inability to control both what they heard over the headset and who they were speaking to via the headset. Because voice is the only channel on *Xbox Live* in which players can converse with other players, the rejection of the voice headset removed the participants' ability to engage with other players in conversation. When participants did try to use the headset, they faced the problem of identifying who was talking and who they were talking to. Participants could not link conversations with actual people and thus were not able to connect socially with the other players on *Xbox Live*.

The inability to ascertain important social information such as the identity of the person who is speaking reflects a lack of translucence on *Xbox Live*. This absence of social cues created disorder and chaos in the voice channel, illustrated by the amount of noise in the channel.

6.2 Dual Social Spaces

The research design created a situation where participants were part of two concurrent social spaces: the group of participants in the laboratory, and the players on the network. We found that participants coexisted in these two social spaces, and were able to fluidly shift their attention from one to the other through suitable use of the voice headset. However, although the voice headset did allow participants to coexist in these two social spaces, our participants found it necessary to develop a number of ad hoc strategies to work around limitations in the design of the headset and better manage the interface between these two

social spaces. Use of the voice channel requires the use of the headset, which is a single-user device that does not lend itself to use by a co-located group of people.

As Jarrett *et al.* [13] indicated, many console gamers like to play in small co-located groups around a single console. In order to incorporate this style of play into *Xbox Live*, it may be necessary to place the microphone and/or loudspeaker more centrally in the room, rather than in a single-user headset.

In the games we used, the audio track was split into a channel for sound effects and music, played through television speakers, and a channel for voice, played through the headset. However users of (non-networked) consoles commonly listen to the game sounds through headphones, to avoid annoying others with excessive household noise. It would be difficult for an *Xbox Live* user to wear headphones as well as an *Xbox Live* headset.

Allowing the user to choose whether voice and soundtrack are separate or intermingled in the same speaker output could solve these problems of distributing sound among social spaces.

Rapid shifting of a user's attention between networked and local social spaces is more comprehensible to players co-located with the user, than it is to fellow players who are present virtually via the network, because a user is more visible to co-located players than networked players. Thus, co-located players can use visual cues to tell where attention and conversation are directed, whereas on-line players must infer where another player's attention is directed from what the other player is saying. This difficulty could be lessened by some means of representing to which social space attention is directed, or signaling who is being addressed during speech.

6.3 Early Evidence of Level 2 Appropriation by Participants

Even though our gameplay sessions were limited to an hour, there were signs that participants were already adapting the technology-as-designed to circumvent limitations and use the technology in novel ways. For example, as previously discussed, users worked around the limitations of the single-user headset in dealing with dual social spaces.

In some *MotoGP* sessions, players chose not to advance the game beyond the lobby screen, in order to continue conversations that had started there. In a similar fashion, one group of participants engaged in extensive conversations with another player during a game of *Unreal Championship*. Gameplay became secondary to the conversation.

A number of participants quickly engaged in exploration of the game world and experimentation with the actions they could take in the world and the effects their actions could have. These forms of activity are the beginnings of evaluation and adaptation of the technology through use.

In these ways participants showed early signs of level 2 appropriation, despite the short duration of exposure to the technology.

6.4 Limitations of the Study

Xbox Live is new. We were limited in the choice of games we could offer participants. It could be argued that the games used were not the most apt to encourage voice communication and social interaction. Many participants felt that *MotoGP*, in particular, was not well suited for voice communication as there is no teamwork needed in the game. The fast paced and competitive nature of the games used in the study may have limited the social interactions that we observed. Social interactions are more prominent in multiplayer RPGs, and voice

may be more suitable to this genre of game. However, it has been suggested that voice communication will detract from immersion in RPGs [2]. The genre of games in which voice will be useful and used heavily remains an open question at this early stage in the life of *Xbox Live*.

In addition, the short amount of time (approximately one hour) that participants had in which to engage with the game necessarily limits our findings. Some participants were unfamiliar with the hand controls used to play the games: a few expended considerable effort to learn the controls, and this interfered with their exploration of the voice channel. Some felt that they needed more time to become comfortable with the interface and to develop sufficient familiarity with other people online to be able to comfortably engage with them in conversation. We were asking participants to play against strangers over the *Xbox Live* network. People's reluctance to engage in conversation with people they do not know may have limited the scope for use of voice communication in our study.

7. FURTHER RESEARCH

Online console gaming, such as that enabled by *Xbox Live*, is in its infancy. The next few years will offer enormous opportunity to study the adoption and appropriation of this new technology.

The study reported in this paper has examined the attractors and repellers associated with the voice communication medium as it is currently implemented in two popular *Xbox Live* games. This represents a study of the first stage of technology adoption as outlined in the MTA. We intend to continue this study and observe how users adopt and appropriate the use of voice communication technology in subsequent stages of the model. This will allow us to gain a deeper understanding of what is appealing, useful and useable about voice communication in multiplayer online video gaming.

A laboratory setting was suitable for gauging users' initial reaction to this new technology. In order to study 'technology-in-use', a more naturalistic setting is desirable. Contextual interviews [3] with players and ethnographically informed observations [4] of game play in the settings where it naturally occurs, such as lounge-rooms and LAN cafes, should offer valuable insights into the adoption and appropriation (or otherwise) of the voice channel in online multiplayer games.

Our study indicates that further work is required to determine how best to implement voice communication in online multiplayer video games. In pursuing this line of enquiry it will be necessary to bear in mind that different genres of games and different styles of game play may have different requirements in this regard. We are aware that researchers at Microsoft are working to improve the implementation of voice in *Xbox Live* games [26].

8. CONCLUSION

The research reported in this paper represents the first stage of a longer study into the adoption, adaptation and integration of voice communication media into everyday online multiplayer gaming practice. It has provided an early look into how users react to their first encounters with *Xbox Live*, a salient example of Computer Supported Cooperative Play. Our key findings were: (a) the participants' ability and willingness to engage socially with others on *Xbox Live* was hindered by the poor usability of the voice channel; (b) the voice channel's poor usability in the games we tested was due to lack of control over what is sent, and what is received, over the channel; (c) players of online multiplayer games exist in two social spaces, the local

and the virtual, and these two spaces can impinge upon one another via the voice channel; (d) while participants expected voice communication to be an advance over text-based communication, in practice they often rejected it.

Perhaps the most significant finding to emerge from this study is that the current implementation of voice in some *Xbox Live* games is socially 'opaque', in that it is difficult to infer socially relevant information that would aid interaction, such as who is saying what to whom. In order for online videogames to be convivial places for social interaction – virtual 'third places' – we need to design for more socially translucent computer-mediated communication by attending to issues of awareness, visibility and accountability of users.

9. ACKNOWLEDGMENTS

We would like to thank Connor Graham for his helpful advice and poetic turn of phrase, Jennie Carroll for discussion of technology appropriation, Steve Goschnick for supporting our use of the IDEA lab, and the Interaction Design Group at The University of Melbourne for constructive criticism of earlier versions of this work.

10. REFERENCES

- [1] Bannon, L. From Human Factors to Human Actors: The Role of Psychology and Human-Computer Interaction Studies in System Design. In J. Greenbaum and M. Kang (Eds) *Design at Work: Co-operative Design of Computer Systems*. Lawrence Erlbaum, Hillsdale, NJ, 1991, 25-44.
- [2] Bartle, R. Not Yet You Fools! *Game+Girl=Advance*: http://www.gamegirladvance.com/archives/2003/07/28/not_yet_you_fools.html <Last accessed 09/12/2003>.
- [3] Beyer, H. and Holtzblatt, K. *Contextual Design: Defining Customer-Centered Systems*. Morgan Kaufmann, San Francisco, CA, 1998.
- [4] Blomberg, J., Burrell, M., & Guest, G. An Ethnographic Approach to Design. In J.A. Jacko & A. Sears (Eds) *The Human-Computer Interaction Handbook*. Lawrence Erlbaum, Mahwah, NJ, 2003, 964-986.
- [5] Carroll, J., Howard, S., Peck, J., Murphy, J. 'No' to a Free Mobile: When Adoption is Not Enough. In *Proceedings of the Thirteenth Australian Conference on Information Systems (ACIS'2002)*, (Melbourne, 4-6 December 2002), 2, 533-546.
- [6] Carroll, J., Howard, S., Peck, J., Murphy, J. A Field Study of Perceptions and Use of Mobile Telephones by 16 to 22 Year Olds. *The Journal of Information Technology Theory and Application*, 4,2, 49-62, 2002.
- [7] Carroll, J., Howard, S., Peck, J. and Murphy, J. From Adoption to Use: The Process of Appropriating a Mobile Phone, *Australian Journal of Information Systems*, 10, 2, 38-48, 2003.
- [8] Czinkota, M. (Ed.). *Marketing: Best Practices*. Dryden Press, Fort Worth, 2000.
- [9] Daft, R. and Lengel, R. Organizational Information Requirements, Media Richness and Structural Design. *Management Science*, 32, 5, 554-571, 1986.
- [10] Ellis, C. A., Gibbs, S. J., Rein G. L. Groupware – Some Issues and Experiences. *Communications of the ACM*. 34, 1 (January 1991), 38-58.
- [11] Erickson, T. & Kellogg, W. A. Social Translucence – An Approach to Designing Systems that Support Social

- Processes. *ACM Transactions on Computer-Human Interaction*, 7, 1 (March 2000), 59-83
- [12] Howard, S., Carroll, J., Vetere, F., Murphy, J. & Peck, J. Young People, Mobile Technology and the Task Artefact Cycle. In *Proceedings of the Eleventh Australia Conference on Computer Human Interaction (OzCHI'2001)*, (Perth, 20-22 November 2001), Edith Cowan University, 13-18.
- [13] Jarett A., Estanislao, C., Dunin, E., MacLean, J., Robbins, B., Rohrl, D., Welch, J., Valadares, J. *IGDA Online Games White Paper 2nd Edition*. International Game Developers Association, http://www.igda.org/online/IGDA_Online_Games_Whitepaper_2003.pdf 2003.
- [14] Kennedy, S. Using Video in the BNR Usability Lab. *Association for Computing Machinery Special Interest Group on Computer-Human Interaction Bulletin*, 21, 2, 92-95, 1989.
- [15] Kutti, K. Activity Theory as a Potential Framework for Human-Computer Interaction Research. In B. Nardi (Ed.) *Context and Consciousness*. MIT Press, London, 1996, 17-44.
- [16] Krueger R. *Focus Groups: A Practical Guide for Applied Research*. Sage, Thousand Oaks, CA, 1994.
- [17] Lawson, R., Tidwell P., Rainbird P., Loudon D., Della Bitta A. *Consumer Behaviour in Australia and New Zealand*. McGraw-Hill, 1999
- [18] Manninen, T. Virtual Team Interactions in Networked Multimedia Games – Case: “Counter-Strike” – Multi-player 3D Action Game. In *Proceedings of The 4th Annual International Workshop on Presence (PRESENCE2001)*, (Philadelphia, 21-23 May 2001) 2001 Temple University
- [19] Multimedia Victoria. *Game Plan: Game On – A Blueprint for Growing the Victorian Computer Game Industry*. Victorian Government., Australia, 2003.
- [20] Ngwenyama, O. and Lee, A. Communication Richness in Electronic Mail: Critical Social Theory and the Contextuality of Meaning. *MIS Quarterly*, 21, 2, 145-167, 1997.
- [21] Nielsen, J. *Usability Engineering*. Academic Press, Boston, 1993.
- [22] Oldenburg, R. *The Great Good Place: Cafes, Coffee Shops, Bookstores, Bars, Hair Salons and other Hangouts at the Heart of a Community*. Paragon House, New York, 1989.
- [23] O'Malley, C. E., Draper, S. W., Riley, M. S. Constructive Interaction: A Method for Studying Human-Computer Interaction. In *Proceedings of IFIP Interact'84*, (London, 4-7 September 1984) 1984 (269-274).
- [24] Preece, J. & Maloney-Krichmar, D. Online Communities: Focusing on Sociability and Usability. In J.A. Jacko & A. Sears (Eds) *The Human-Computer Interaction Handbook*. Lawrence Erlbaum Mahwah, NJ, 2003, 596-620.
- [25] Rheingold, H. *The Virtual Community: Homesteading on the Electronic Frontier*. Addison-Wesley, Reading MA, 1993.
- [26] Terrano, M. Lessons from Life: Designing More Immersive Games. Paper presented at the Australian Game Developers Conference (Melbourne, 20-23 November) 2003.
- [27] Turkle, S. *Life on the Screen: Identity in the Age of the Internet*. Simon & Schuster, New York, 1995.
- [28] Wadley, G., Gibbs, M., Hew, K., Graham, C. Computer Supported Cooperative Play, “Third Places” and Online Videogames. In *Proceedings of the Thirteenth Australian Conference on Computer Human Interaction (OzChi 03)*, (Brisbane, 26-28 November 2003), University of Queensland, 2003.
- [29] Wright, T., Boria, E., Breidenbach, P. Creative Player Actions in FPS Online Video Games: Playing Counter-Strike. *Game Studies*, 2, 2, (December 2002), <http://www.gamestudies.org/0202/wright/>