COPYRIGHT

11.FROM TEXT TO TALK:
MULTIPLAYER GAMES AND VOICEOVER IP

John Halloran
Yvonne Rogers
Geraldine Fitzpatrick

ABSTRACT

The social experience of multiplayer gaming is mediated by the communications tools that are available to use. Until recently, these have been largely text-based, but with the advent of new voiceover IP tools like Roger Wilco and Xbox Live, voice-mediated communication is becoming increasingly common. We present three studies of multiplayer gaming, where we analyse what happens in terms of the social experience when players are given the opportunity of talking to each other rather than texting. To do this we use a conceptual framework called FFIPS, which stands for Form, language Functions, Identity, Presence, and Social protocols. Our findings show that voiceover IP for multiplayer gaming appears to be well-suited to supporting a distinctive and enjoyable social experience, both by providing high 'presence' (i.e., increased energy, engagement and vividness), and by revealing information about players' real identities.

INTRODUCTION

The scaling-up of computer games from single-player to multiplayer has meant that gaming can now offer an experience that is quite different from that of playing alone against a computer. Rather than an individual pursuit, multiplayer gaming has become a social experience - one that can constitute a 'social pleasure' [7].

The nature of the social experience of multiplayer gaming depends on a number of factors, including connection speed, whether players are co-located or distributed, how many people are involved, who they are, and whether they friends or strangers. The type of game is also important. There are many, ranging from 'role play games' (RPGs), where players can become characters on a quest, through 'first person shooters' (FPSs), which involve fighting against opponents in war settings, to race games, and beyond.

The majority of contemporary games feature detailed, realistic 3D virtual worlds that players navigate through as an avatar. Within these worlds, depending on the type of game, players need to communicate with each other

for several reasons including discussing strategy, calling for help, commenting on performance, or just chatting. Until recently, communications were text-based. In RPGs, conversation appeared as text boxes above avatars' heads. In FPSs, scrolling text strings would allow one player to do things like congratulate another on a kill, or give information concerning their location.

With the integration of voiceover IP into computer gaming, players can now use tools like Roger Wilco and Xbox Live to talk to each other, making text communications unnecessary. This new development has left games producers eager to sell games not just for entertainment value, but for their potential to enable players to interact with friends in new ways, meet new people, and even form new relationships [15, 16]. The aim of this paper is to gauge the extent of that potential by examining how the ability to talk, rather than text, affects the social experience of multiplayer gaming.

To make our analysis, we use a conceptual framework called FFIPS, which stands for Form, language Functions, Identity, Presence, and Social protocols. This is a set of concepts we have found key in our research into how different types of communications tool affect the activities they support. The purpose of the FFIPS framework is allow us to compare different kinds of talking in different kinds of game settings, and to relate this research to the wider context of computer-mediated communication and collaboration.

TEXT- AND VOICE-BASED COMMUNICATIONS TOOLS: AN OVERVIEW

Research into the issue of how voiceover IP impacts the social experience of multiplayer gaming, and how this might differ from using text, has only just begun. Here, we review relevant related research in CMC (computer-mediated communica-

tion), MUDs (multi user domains), and CVEs (collaborative virtual environments) – as well as looking at recent work on text communication in multiplayer games.

CMC and text-based communications

Computer-mediated-communication tools, which have been largely text-based, have attracted much attention over a number of years. A significant research question has been how these can affect and change the nature of communicating and socialising. An important concept in CMC is 'social presence'. This concept originally referred to the notion that communications that are not face-to-face cause 'psychological distance', which reduces sociability [9]. More recent work argues that social presence is the degree to which a human actor can be perceived through CMC [3, 5]. All these approaches share the assumption that CMC involves attenuation of the communicative resources available in face-to-face interaction, with effects on the social experience.

An early extension of social presence, reduced social cues ('RSC') theory [10] argues that in communicating face-to-face we make use of a number of verbal and non-verbal social cues. In CMC, there is an absence of visually transmitted social cues, and this can lead to 'disinhibited' behaviour, for example e-mail 'flaming' where people exhibit greater aggression or frankness than they would face-to-face. According to RSC, the reason this happens is that because interlocutors are not visually present to each other, they are more self-oriented and less aware of others. This raises concerns about how people may misrepresent themselves through online behaviour using CMC with possibly detrimental results.

In contrast, more recent research suggests that reduced social cues can have marked positive effects on the social experience of CMC [13]. The physical

absence of the interlocutor can lead to reduced self-presentation concerns that can allow people to more easily self-disclose, and others to reciprocate. This can set up positive feedback loops, where intimacy rather than hostility occurs.

These findings suggest that text-based CMC can have a direct impact on the social experience of interacting with others online. The anonymity, increased awareness of self, and reduced awareness of others associated with CMC can change how people communicate. In particular, CMC can lead to alterations in identity compared with face-to-face interaction, not only in regard to how people present themselves, but also how they perceive others.

Text-based communications in MUDs

Writers like Turkle [12] and Reid [8] raise a new question for CMC: what happens when people, rather than presenting themselves in different ways than they would in face-to-face situations, make use of the properties of CMC to create radically new identities? Their research into text-based MUDs has important implications for social experience. According to this, people can create parallel identities that enable them to construct and experiment with sexuality, race, gender and power. These identities may be validated online in ways which make the social experience powerfully attractive. However, here, the construction of identity becomes less an artefact of the attenuation of cues in face-to-face communication, and more a complete departure from what might hold in face-to-face 'reality'.

Voice-based communications in CVEs

In recent years there has been much research into collaborative virtual environments (CVEs). These are three-dimensional virtual worlds that can be used for a variety of purposes including collaborative performance, meetings, and work. They often include avatars to represent participants, and can feature voice-based communications tools. Research in this area might help us understand how voice-based communications tools work in virtual worlds, an issue very relevant to multiplayer gaming.

Bowers *et al* [2] look at how talk and embodiment function in CVEs designed for meetings. They note the problem of discontinuity between avatars and voice-mediated communications. According to this research, people find it hard to take turns, preferring to wait for others. Embodiments can be used as ways of signalling to others that they may speak (for example, through turning and facing), but can tend not to be, so that verbal means like 'scanning' – using talk to find out who is online and who wants to speak – are required. This reflects a lack of coupling between avatar actions and verbal actions.

Studies of work mediated by CVEs help reveal what sort of coupling between the virtual world and people's spoken interaction needs to occur. For example, Tromp's HTA (hierarchical task analysis) [11] aims to uncover what kinds of generic tasks collaborators need to be able to carry out in work settings. These include turn-taking, shifts in avatar proximity, shifts in avatars' relations to artefacts including virtual documents, and indexicality (the ability for the avatar to point something out and refer to it using context-dependent cues like 'here', 'there', 'that'). Such research implies that for an effective social experience to take place, talk needs to be integrated with avatar actions in ways which can restore the postural, gestural and proximity information that embodiment provides. Talk is also needed to help mediate collaborative performance art. This can include the interaction of real people with avatars in virtual spaces [1], which requires a high level of 'orchestration': the interaction presents levels of

challenge which make a production crew and assistants necessary.

This research on various types of CVE shows that integration between task, visual representations (including avatars, documents, furniture etc.), and communications tool is challenging. How the communications tool works (or does not work) is strongly related to these other factors. Thus, we might expect to see interdependencies in multiplayer games, too.

Text-based communications in multiplayer games

The research discussed so far does not look directly at the social experience of computer games. However, it suggests several ways to look at text-based communications in computer gaming, and ways it might compare with voice-based communications. Questions that arise include: Does text mediation in computer games lead people to present themselves in different ways than they would face-to-face, or does it allow the creation of radically new identities? How does a specific type of game affect it?

Recent research into text messaging in FPSs [6, 14] has started to reveal innovative types of talk particular to this gaming context. These include creating new kinds of alias such as 'Smoke Weed and Kill People'; 'Mark Killer'; 'Osama Yo Mama', and so on. This is evidence of identity management which, rather than departing from reality, can engage current social concerns in ways designed to shock others (e.g. drugs, crime, terrorism). However, this research shows that other types of behaviour occur which contradict the notion that FPS players want to transgress social norms. Much talk is highly skilled, concerning the giving and eliciting of tactical information, elicitation of levels of expertise of other players, discussion of technical issues like lag

(whether there is delay in graphics display), and even 'policing', whereby gamers that transgress gaming etiquette are rejected or 'kicked off'. Other talk is 'creative', including joking and irony, collaborative rule-changing, popular culture references, or 'performance talk', concerned with things like greeting, discussing strategy, congratulating, scorning and so on. This research shows that, notwithstanding the creation of novel aliases, text-based communications in FPSs are often directly connected to players' actual levels of expertise and experience.

This short overview reflects that text-based communications in computer games are highly developed as well as variegated. The forms of communication that take place seem to depart from the issues we identified that affect both CMC and MUDs. Texting in FPSs does not appear to lead to exaggeration of hostility or intimacy, and appears to be associated with more modest identity creation and experimentation than can happen in MUDs.

THE FFIPS CONCEPTUAL FRAMEWORK

Our literature overview reveals several issues related to different communications tools when used for a variety applications and activities, with different effects on the social experience. These issues provided the basis from which we developed our FFIPS conceptual framework (see Figure 1).
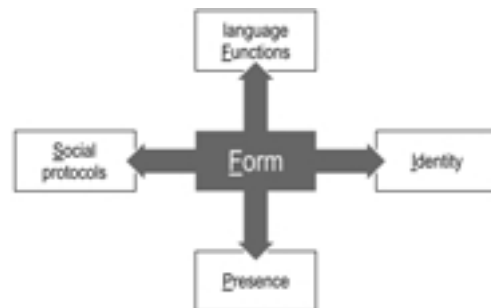

Figure 1: The FFIPS Conceptual Framework

Form is at the centre of the FFIPS framework. This concept is used to define (a) the type of communications resource, for example a text messaging interface or a voiceover IP tool; and (b) its context. Is it, for example, part of a MUD supporting an RPG, or of a CVE to support collaborative performance? We also use four other concepts: language functions, identity, presence, and social protocols. All of these are influenced by form, reflected by the arrows. What are these concepts and why are they important?

Language functions include, for example, greeting, persuading, supporting, etc. This concept relates to how people get things done socially by means of talk. Its use in the framework is to help identify where a communications resource enables or disenables this; and how the social experience is affected.

Identity is an important issue across much research concerning communications tools. In FFIPS it is used to consider how the social experience of is affected by how far a communications tool allows identity to be exaggerated, managed, created, or perceived.

In the FFIPS framework, presence has a specialised use. It refers, like the concept of social presence, to how far social cues are preserved by the communications resource. In addition, it is used to consider how far a communications resource contributes to immersion in a convincing virtual world, and also to how vivid, energised and engaging the social experience of that world is.

The remaining concept is social protocols, which considers the issue of how people go about negotiating social episodes using communications resources, and what are the rules and procedures involved. The concept is used to help decide whether there are shared understandings of how to behave socially, how far communications resources support this, and what are the effects on the social experience.

Throughout the rest of the paper we exemplify the FFIPS framework and show how it can be used to draw out how the social experience can change when people are able to talk instead of texting in multiplayer games.

FROM TEXT TO TALK: THREE STUDIES OF COMMUNICATIONS TOOLS IN MULTIPLAYER GAMING

We carried out a series of studies to explore how groups of players socialize when gaming, and in particular what types of talk they use. We were also interested to see if they change the way they talk relative to face-to-face interaction, when they talk through voiceover IP tools.

Study One: Eight Halo players in the same room

Our first study aimed to find out how talk is used when multiplayer gamers are able to talk to each other face-to-face. Using the FFIPS conceptual framework, we wanted to see how a range of issues might affect the social experience. One issue was what kind of voice-mediated interpersonal interactions occur. Another was what interactions there are with the visual material presented by the virtual world of the game.

We observed a group of eight experienced multiplayer gamers in their early-to-mid teens over three meets which lasted around an hour each. The gamers had been playing together for over six months on a fortnightly, and occasionally more frequent, basis. At each meet we set up a video camera on a tripod and left the room so as not to interrupt the flow of interaction. Our analyses are based on the resulting video data.

This group favoured Halo, a fast-moving, exciting FPS played on Xbox consoles over a LAN. The larger

group split into two sub-groups who, using two separate TVs, played against each other in the same room. Each game lasted a maximum of 30 minutes. There would be 'mixing and matching' among the eight so that the two teams were constantly changing membership from one game to the next.

A consistent finding across all three sessions was that there was a great deal of simultaneous talk, with gameplayers shouting and talking across each other in a loud and at times chaotic way. Another finding was that utterances could be reduced to a limited range of language functions. These were (1) 'joshing': jokes or irony, e.g. 'man you are so SICK'; (2) 'crowing': celebrating one's own achievements, those of another, or their misfortunes e.g. 'Ha ha you're DEAD!'; 'NICE kill!'; (3) strategy talk: e.g. 'I need a gunner'; and (4) side- or self-talk e.g. 'Oh that was SO rubbish...'. The verbal behaviour we saw was associated with a lot of laughter and physical movement (leaning forward, leaning back, shifting, 'punching' the console). We also saw other events which were non-verbal, but afforded by co-location - like the simultaneous arm-raising and cheering by the winning team shown in Figure 1(a); and the rapid reorganisation shown in Figure 1(b).

Figure 2: Halo Players: (a) whooping/handslapping; (b) reorganising

In terms of our FFIPS conceptual framework, the form of the communications resource was face-to-face talk, in the context of co-located Xbox console gaming using Halo. This was associated with different language functions than are found in CVEs. This may be because different social protocols hold. In CVEs for meetings, it is important for people to take turns, not to talk over each other, and to make clear who is being addressed. This is the opposite of what was allowed, and apparently encouraged, by the Halo gaming.

Another reason for the simultaneous talk and the different kinds of language functions may be that the utterances, although coupled to game events, were not necessary to achieve the performance of the game in the way that verbal communications in CVEs often are. Only strategy talk is important in this respect, but we saw unexpectedly little. It appears that language functions in this study did not need to relate to problem-solving as much as for a CVE because an FPS as a task is well-known and often repeated by experienced gamers like those we observed. In this context, language functions associated with joking and having fun were much more in evidence, as were associated social protocols which allowed loud simultaneous talk without specific addressees.

Communications in co-located multiplayer gaming (also known as 'LAN parties') are face-to-face, and this places constraints on how far identity can be manipulated. There were two kinds of interaction: (1) the interaction of avatars with other avatars in the virtual world of Halo; and (2) the interactions between the players in the room, which were both verbal and physical. These parallel interactions blur the disjunct between player and avatar and suggest that identity might be more continuous in co-located contexts, than in distributed contexts where the user associated with an avatar cannot be so readily perceived by others. While utterances like 'I need a gunner' show the players taking on game-associated roles, their 'real' identities were known to each other.

The roleplay appears to have been part and parcel of the experience of playing a game with friends, rather than being an example of the kind of identity manipulation which can happen in MUDs.

This type of gaming also has implications for presence. As we have seen, 'presence' can refer to social presence, and to immersion in a virtual world that seems real. In the Halo gaming, it might be expected that the amount of noise and activity in the room could have distracted the players from immersion in the virtual world, but also, that the virtual world might have meant players were less able to attend to the social presence of others. However, it appears that the opposite happened: each experience amplified the other. A LAN party appears to create a special form of presence, one that is highly engaging with high energy levels - a vivid, 'live' event.

The social experience of co-located multiplayer gaming is, on this evidence, highly energised and enjoyable, with a limited range of language functions, and social protocols which encourage simultaneous talk. This sort of gaming involves a coupling of the virtual world of the game to the real world of the room, which results in high levels of presence.

Study Two: A singleton Xbox Live player

In contrast to the Halo study, which examined talk in co-located gaming with friends, our second study aimed to look at how geographically distributed gamers talk online to people they do not know.

We observed a singleton player, Joe, 23, over two Xbox Live sessions of an hour each ('Joe' is not the participant's real name). Xbox Live gaming consists of an Xbox console through which players can select other players online, plus the Xbox Live headset which plugs into the console and allows players to talk to each other. During each session, Joe played three of his favourite games, 'Unreal Tournament' (an

FPS); MotoGP (a race game); and 'Whacked' (a 'tag' game where players find and hit each other with a range of implements). We video-recorded each session, and also asked questions during the gameplay.

A finding that held across the two sessions was that talk was much quieter than in the Halo study. Joe's tone of voice was even and measured, with a 'bland' feel. Utterances were less frequent, but with a greater number of language functions. However, there was some decoupling between talk and gaming: what was said frequently bore little relation to events in the game. Another finding was that, although Joe appeared to be enjoying the experience, energy levels seemed lower than in the Halo gaming. We also found, even though Joe's identity was revealed to others only through the virtual world of the game plus his voice over Xbox Live, that he did not attempt to manipulate his identity. He also engaged in apparently formulaic ways of talking which suggest that there are well-understood social protocols for Xbox Live gaming.

In terms of FFIPS, the form of the communications resource in this study was voiceover IP in the context of Xbox Live gaming with strangers. This was associated with three of the language functions we saw in the Halo gaming – joshing, crowing and self-/side-talk – but no strategy talk, as this player was not involved in a team effort. In addition to these, other language functions occurred: 'scanning', greeting, and 'scoping'. By 'scanning' we mean that Joe searched for other users by repeatedly saying 'Hello? Hello? Anybody there?'. This utterance, which did not vary in its form, served three purposes: (1) to see who else was online; (2) to start to talk to others he could see were online; and (3) to establish whether he was able to talk to others at all - in MotoGP, for example, the player can only talk to the racer in front and behind, to free up CPU time for

graphics. In terms of greeting, when someone talked to Joe for the first time, Joe consistently used the same formulae: 'How you doing mate?' for a male player, and 'How you doing?' for a female. These did not vary. Having scanned and greeted, Joe would do some 'scoping'. By this we mean questions concerning nationality, age and so on to glean information. Example utterances included 'Where you from?'; 'Are you American... Canadian... from Montreal?'.

The formulaic utterances have implications for identity. One big difference from the Halo study is that the players did not know each other before playing together. The initial language functions we observed – scanning, greeting and scoping – may be formulaic (a) to allow formulaic responses, which may be socially easier; and (b) to protect identity until more information is known about another player. These language functions, related to the form of communications (voiceover IP, implying geographically distributed players), appear to have little to do with identity effects like hostility or intimacy. Rather, they seem to be associated with establishing a bland, non-committal form of initial self-presentation. This blandness was supported by the measured, low-volume speech which accompanied these functions. However, while identity appeared to be protected, and may be subject to the principle of reciprocity of self-disclosure, this was not associated, in this study, with the creation of false identities. Joe told the truth about who he was, where he was from, what games he liked and how long he had been playing them; and, as far as we could tell, so did his online interlocutors. This suggests that voice-mediated communications might not be associated with identity manipulation and management in the same way as text-based social experiences have been.

We have already noted that communications, while they may not be necessary in order to perform a game, can be coupled to it. However, the observations of Joe often showed a lack of coupling as if he were engaging in two simultaneous, but different, social experiences. One of these consisted of chasing an avatar in Whacked and repeatedly hitting it; the other of a conversation with the controller of the avatar, to establish where that person was from, how long they had been playing, and how old they were. This reflects a different kind of presence from that found in the Halo study. On one hand, Joe appeared to want to create social presence through his questions, since the player could not be seen. On the other, this may have reduced immersion in the virtual world. This reflects that verbal communications in multiplayer contribute to the social experience in ways that can differ, depending on whether gaming is co-located or distributed, and whether people know each other.

Study Three: Three soldiers of fortune

The purpose of this study was to look at distributed multiplayer gaming supported by voiceover IP when the players know each other, as opposed to being strangers, as in the Xbox Live study, and to see how this might differ to co-located multiplayer gaming (the Halo study). This study also gave us the opportunity to compare the social experience of the same game, supported by talk, or by text only.

We identified three experienced players – Saleh (21), Chris (22), and Zak (21) (these are not the participants' real names) – who used PCs rather than consoles. The three lived at the same address with a PC in each of their (separate) bedrooms. They formed a clan who had been playing together for 6 months 'several times a week'. Their favourite game was Soldiers of Fortune, an FPS. The group claimed that they rarely played any other. The clan had made its own modification to this gaming experience by adding Roger Wilco so that they could talk to each

other, unheard by others they were playing against. We video-recorded this group over four sessions of 60 minutes each. In two of these, the gamers used text only, using the tools provided by Soldiers of Fortune. In the remaining two, we asked the players to use talk rather than text, and in addition to video, we recorded the audio conference. We also carried out two participant analysis sessions with the group, playing back recordings and asking open-ended questions about gameplay which was too fast moving to be susceptible to questions at the time.

Our findings for this study differ from the other two (Halo co-located gaming and Xbox gaming with strangers). Where the gamers used talk, we found that, while there were similar language functions to the Halo study, there were fewer utterances. There could also be long stretches of silence. Utterances tended to be made at low volume, but although talk was quieter and more intermittent than in the Halo study, the players seemed deeply immersed in what they were doing. We also observed that

the gameplayers' talk appeared to be based on a good deal of implicit knowledge, both about the way the game worked, and of each other. As we will see, these findings have implications for identity and social protocols.

Where the gamers used text, their gameplay as a clan was less cohesive. Zak, Chris and Saleh had more trouble coordinating strategy and their scores were lower. The group produced virtually no text messages, and during the participant analysis they explained that text communications for Soldiers or Fortune was something they now dislike, much preferring the social experience of being able to talk.

FFIPS predicts that the form of a communications resource will affect language functions, identity, presence (as defined in the framework), and social protocols. All these influences are shown in the following excerpt from a talk-based session (Figure 3), which lasts around 45 seconds. The pictures show Saleh, whose utterances are prefixed 'S'. The only

S(02): Yo dese guys is jokers man

S(04): I shot 'im in de head and he's still like alive shootin' at me

Z(09) Chris where did he shoot you

S(33) Yo guys I'm stuck in some room up 'ere yeah

S(35): De player's shootin' at me

Z(46) Automatic shotgun

Z(53) I see Saleh

S(56) AARGH

Figure 3: Playing Soldiers of Fortune with Roger Wilco

other speaker is Zak ('Z'), although Chris is online. The numbers represent the time, in seconds (starting from zero) where the utterance commenced.

Chris and Zak tended to lead all the gaming sessions in terms of 'kill rates' and strategy. A good deal of their success related to their experience of the game. Both were highly familiar with 'maps', i.e. the virtual architecture of the game (usually a large building like a hospital or hotel). A major aim for these two was to make sure they knew where each other was, in relation to other team members and to the opposite team, to coordinate attacks, but also retreats. They also needed to let each other have information about their 'health' (i.e., of the number of lives allowed for each game, and how many each had left), and what weapons they had at their disposal.

In the excerpt above, Chris does not communicate verbally at all – and in general, he spoke the least. Zak speaks three times asking Chris first 'where did he shoot you', asking for confirmation of Chris's location when he was last killed; 'automatic shotgun', announcing he now has this weapon; and 'I see Saleh', announcing to Chris that the remaining team member has been found. What is striking here is Saleh's apparently unsuccessful attempts to engage the other players in his own problems. Saleh is under fire, unable to say where he is exactly – 'Yo guys I'm stuck in some room up 'ere yeah' – due to less knowledge of the maps. He also asks for help when he is under fire. None of this gets a response. However, Saleh may not expect to be acknowledged. He tended to assume an argot when playing the game, an exaggeration of his normal speech, which suggests he is attempting to join a club (his clan), but also that he is simply assuming an enjoyable, and humourous, role. Thus there appears to be an implicit understanding that Zak and Chris will communicate and remain aware of each other, that when they speak

they are addressing each other; and that Saleh will be left to his own devices to coordinate his actions with the other two.

This analysis shows that there is a coupling, as in some CVEs, between task structure and voice-based communications. Here, it is mediated by implicit knowledge not only of the game but also of social relationships. The players each know what their relative level of skill is, and this is reflected in social protocols which allow utterances to be successfully addressed to other players without the addressee being made explicit. Equally, there appeared to be shared understandings that verbal responses may not be required. These social protocols also affect identity: Saleh's identity as a (comparative) learner appears to be reinforced, while Zak and Chris preserve their identities as experienced leaders. This kind of identity management is not concerned with making use of the properties of a communications tool to present in a particular way; or with creating alternative identities. Rather, the effect on identity of the Roger Wilco add-on is to enable the players to project themselves according to a shared understanding of their place in a team. This differs from the Halo gaming in that, although everyone can hear everyone else, utterances are measured and tend not to overlap, and are integrated with the virtual world of the game rather than creating a social experience in the room around the game.

Asked about playing the text version of this game, Saleh observed, 'I felt a bit weird really, playing it, as if I was missing something crucial'. Pressed on whether there was a difference between playing Soldiers of Fortune with text-based versus voice-based communications tools, Zak said, 'I dunno, I find it restrictive. As in tactically, as in gameplay-wise, you're not as secure as with voice. I can just say Chris where are you, Saleh where are you, and I know

that they're on their way almost, like I can wait there, and if you can't talk to other team members, and typing messages in, you're not going to be typing in messages anyway'. This reflects that none of these players were willing to engage in the ergonomics required to text message (done with the left hand) whilst navigating (done with the right). Chris said: 'I don't think you can really compare the two. There is such a big difference with the voice. I just have to say 'where are you guys' and in a few seconds I will get the reply. In the other one I would have to type it, and whilst I'm typing it I can't defend myself'. When asked whether needing to do this kind of thing might lead to a more intense or enjoyable gaming experience due to the added challenge, the three players appeared nonplussed: Saleh said: 'I agree with the statement that it's more difficult. I personally don't enjoy it more. It was really feeling weird. If you can't talk you don't know where they are, you don't know if you can keep in touch with them. I wouldn't play it as much if we didn't have the voice thing.' What this suggests is that the mutual awareness afforded by voice-mediated communications is crucial to the social experience for these three players, this being linked with a high priority for them: strategy. The support for this form of awareness provided by Roger Wilco means that the experience of using text-based communications cannot compare, despite lower perceived level of challenge.

This study suggests that there is an important interaction between knowledge of the game, experience of gaming with known others, and the communications tool associated with the game. This has implications for mutual awareness and attention, and the mutual interpretation of the meaning of utterances. In terms of the social experience, while talk was more intermittent, lower volume, and less frequent than in the other two studies, there was still a very high degree of presence, although of a different type to the Halo study.

This suggests that voice-mediated communications, when used by a clan over a period, can lead to the members experiencing the game, as well as each other, in more engaging and intense ways.

DISCUSSION & CONCLUSIONS

In this paper we have used a conceptual framework we call FFIPS to start to scope out issues and directions which may be important when considering the transition from text to talk in the mediation of multiplayer games. The framework allowed us to analyse the social experience of different gaming contexts in terms of key concepts, and to compare how social experiences differ given different types of communications resource.

In our discussion of CVEs, we identified an important issue to do with coupling: it is necessary for talk to be coupled with a virtual world, particularly avatar actions, in ways which are essential for a satisfactory social experience to occur. However, in all three studies, players had no problem in acting in the virtual world of the game supported by talk. This reflects that where actions and events in virtual worlds are well-known and familiar, talk does not need to bear a cognitive load related to problem solving. Rather, it can be freed up to serve a range of functions relating to fun and enjoyment.

Talk appears well-suited to supporting the social experience of multiplayer gaming in ways that go beyond text. There are differences between the two media [4]: talk is immediate, and speakers know that an audience has heard. Successful talk implies 'grounding', whereby there is shared understanding. In contrast, text requires tools, may not be picked up by the audience, and may not be 'grounded'.

The properties of talk mean that where players know each other, there are high levels of presence. Players

seemed engaged and immersed in the social experience, whether it involved loud simultaneous talk or quieter, less frequent utterances. Being able to talk appears to influence presence in important ways. In co-located gaming, it helps couple a virtual game world to a real experience happening in the surrounding room. In distributed gaming, it allows superior gameplay, which leads to greater immersion in the virtual world.

Talk also has important implications for identity. A striking finding across all the studies is that identity creation is not a major issue for the types of game discussed; rather, there are various reasons why 'real' identity persists. In co-located settings in particular, it appears to be an important requirement of the social experience that people get to know each other better. This runs counter to literature on CMC and MUDs which shows that identity is altered by text mediation. In our studies, while people liked to take roles, they also appeared to enjoy experiencing people already known to them in new settings, as well as getting to know new people. Voiceover IP appears well-suited to this pro-social process, and this supports the view of the games industry that voiceover IP has social potential.

Our research has implications for designing voice mediation for games. The social experience of multiplayer gaming using talk appears to depend on players' being able to feel confident that everyone can hear everything that is being said. For this reason, players should have as much auditory access to other players as possible. This implies that the balance between CPU time for (a) graphics and (b) voice, which can limit this access, may need to be reconsidered. Where voiceover IP cannot be used due to processing power being switched to graphics, one solution has been to allow players to send each other pre-recorded voice samples. However, the

social experience seen in the Soldiers of Fortune study, where there was improved gameplay and high presence as a result of talk, may not be possible with this type of short-cutting.

ACKNOWLEDGMENTS

REFERENCES

1. Benford, S., Fraser, M., Reynard, G., Koleva, B. & Drozd, A. (2002). Staging and evaluating public performances as an approach to CVE research. *Proc. CVE '02*, 80-87.

2. Bowers, J., Pycock, J. & O'Brien, J. (1996) Talk and embodiment in collaborative virtual environments. *Proc. CHI '96*, 58-65.

3. Bradner, E. & Mark, G. (2002) Why distance matters: effects on cooperation, persuasion and deception. *Proc CSCW '02*, 226-235.

4. Clark, Herbert H. (1996) *Using Language*. Cambridge University Press.

5. Lee, K. M. & Nass, C. (2003) Designing social presence of social actors in human computer interaction. *Proc. CHI '03*, 289-296.

6. Manninen, T. (2001) Virtual team interactions in networked multimedia games. Case: 'Counter-Strike' – multi-player 3D action game. *Proc. PRESENCE '01*.

7. Poole, S. (2000) *Trigger Happy*. Fourth Estate.

8. Reid, E. (1996) Identity and the cyborg body. In Ludlow, P. (Ed.) *High Noon on the Electronic Frontier: Conceptual Issues in Cyberspace*. MIT Press.

9. Short, J., Williams, E. & Christie, B. (1976) *The Social Psychology of Telecommunications*. Wiley.

10. Sproull, L. & Kiesler, S. (1991) *New Ways of Working in the Networked Organization*. MIT Press.

11. Tromp, J., Steed, A. & Wilson, J. R. (2003). Systematic usability evaluation and design issues for collaborative virtual environments. *Presence*, 12:3, 241-267.

12. Turkle, S. (1995) *Life on the Screen: Identity in the Age of the Internet*. Touchstone.

13. Walther, J. B. (1996) Computer-mediated communication: impersonal, interpersonal and hyperpersonal interaction. *Communication Research*, 23, 3-43.

14. Wright, T., Boria, E. & Breidenbach, P. (2000) Creative player actions in FPS online video games. *Game Studies*, International Online Journal of Games Research. Online article accessed at http://www.gamestudies.org/ 0202/wright/ Verified 31.08.03.

15. Everquest invokes dungeons and dragons spirit. Online article and video presentation accessed at http://www.cnn.com/2001/TECH/08/17/ everquest/ Verified 31.08.03.

16. Role-playing fits the bill. Online article accessed at http://www.guardian.co.uk/online/ story/0,3605,889297,00.html Verified 31.08.03.