Cameras in video games:

Comparing play in Counter-Strike and the Doctor Who Adventures.

Eric Laurier (*Edinburgh University*)

Stuart Reeves (*University of Nottingham*)

1. Introduction

3D video games remain a rapidly evolving new medium bringing together gameplay, narrative, architecture, computing and cinema. Correspondingly for a number of existing schools of inquiry they raise absorbing questions about game theory, storytelling, space, human-computer interaction and the moving image. In the space of this chapter our ambitions are necessarily modest, we wish to concentrate on only one of the devices that arises out of the trading between cinema, television and gameplay that is nevertheless at the heart of the video-ness of video games: the camera. How we play video games is dependent on and generated by the uses of the camera: a 'game-camera' that is modelled upon the optical set-ups of mechanical and digital cameras and yet, as we shall see, is distinct from them. For instance, the game-camera is an invisible point without virtual representation. Game-cameras are central to playing 3D video games because they are a mediated view into, and perspective upon, the 'visibility arrangements' of the game (Watson, 1997). They are at the core of analysing, at any moment, what has just happened, what is currently happening in the game and what might happen next. The interplay between using the game camera and the 'visibility arrangements' of the game produce not just the experience of 3D spaces but the very play of the game.

Accordingly, the present chapter explores the relationship between gameplaying and the carefully crafted movements of visual perspective that players employ in and through the 3D environments that are characteristic of modern video games. Much of the methodical 'work' of play engaged in by players involves rendering complex game spaces intelligible through scrutiny of the virtual environment (VE) and linking together sequences of achieved perspectives. We say 'achieved' because even the so-called 'first-person' perspectives that are provided by the movement of the camera are *not* isomorphic with the glances that may be made by the naked eye. Instead, for both first- and third-person perspectives, visual awareness is achieved by looking with, and through, camera lenses (Macbeth, 1999; Mondada, 2003). Thus, competence in the manipulation of video perspectives becomes an essential acquisition for players as they seek to engage with the game itself and / or with other players. Through instructive comparisons with other settings in which cameras are used to 'look around' and render the scene intelligible (e.g., surveillance (Luff et al., 2000) or endoscopic medicine (Mondada, 2003)), the findings of this chapter will help inform a wider understanding of the manner in which environments are made sense of through video technologies.

1.1 First person and third person

Our analysis is based upon a set of vignettes drawn from a small corpus of video recordings of a 'first-person shooter' that, as the name implies, has a first person perspective, and, an interactive, narrative game, which, by contrast, involves third-person perspectives. A characteristic feature of first-person games is that players experience the 3D environment 'through the eyes' of an avatar (i.e., a virtual embodiment). By virtue of this to move the character is usually to move the camera and, of game-relevance, to move the camera is to move the person. However within the games it is sometimes the case that lenses can be attached (e.g. sniper sights, binoculars) to allow zooming in and out. Turning to third-person games they show a view from behind (and usually above) the character. The player's avatar and the avatar's relationship to the course of play within the game are made visible via the game-camera, which has an indirect coupling to avatar movements. Thus, for third-person games, the view in and around the game's environment can typically be manipulated by the player to inspect the scene without actually moving the position of the player's avatar within the environment. Because it is coupled to the avatar the camera cannot move around the game independently of moving the avatar.

1.2. Cameras-in-interaction

Early work in praxeological studies of the film camera by Macbeth (1999) made clear the differences between looking with human eyes and looking with a (mechanical) camera. When a camera moves quickly from a first perspective to a second perspective it both creates a blur and also struggles to find the second perspective. As Macbeth (1999) noted, the quick glance that members of society make from one thing to a second thing and back again, without missing that second thing, are almost impossible with a camera. Glancing is a routine technique for how sighted members of society remain aware of settings and activities within those settings. In all manner of settings members organise their actions for their glance-availability, and make available their monitoring of those settings through the visibility of their glancing (Sudnow, 1972). Glancing's absence from camera operation is specifically significant because managing the visibility of actions of one's own avatar, other players, opponents, or obstacles is central to the skills of playing video games (Reeves, Brown & Laurier, 2009).

Macbeth’s early work on camera movement has been taken up in other settings where cameras are used (Licoppe & Morel, 2012; Lindwall et al., this volume; Mondada, this volume). In endoscopic surgery (Mondada, 2003), surgeons direct camera-operators both to perform the surgery but also to show students what they are doing. As part of this the surgeon makes explicit and tacit requests to the camera operator. The requests precede next appropriate actions, so that when the camera is moved it then provides the visibility of the thing that the surgeon wishes to work on and/or show to their audience of students. The availability of the image is then an ongoing concern of the surgeon as it is, of course, for video game players. The camera operators themselves track the course of action of the surgeon, and their skills are apprehended in how they will routinely provide the appropriate adjustment of the camera without the need for a request by the surgeon. While these paired tasks in surgery might seem tangential to gameplay, as we will see later, having two people (or more) involved in a game raises similar issues around coordinated action.

Taking up Mondada’s approach, Broth (2004, 2009) explored how multiple camera operators in live broadcast, offer up shots to their editing team. Rather than doing this verbally, camera operators make visible their offers by rapidly panning or zooming the camera and then stabilising a shot of a speaker. The brisk movements of the camera, in the context of the production room setting, become see-able as actions such as ‘the operator searching for a shot’. In the course of such a search when the camera settles and has a face in focus in its middle, then at that point, this person is being offered for the editor to select for broadcast. What both Mondada and Broth’s studies bring to the fore is how the movement, framing, zooming and focus of the camera make actions visually available to the parties (e.g., responses, proposals, searches). Framing and zooming perform similar work in the operation of cameras in games.

Finally, of most obvious relevance to our chapter is Keating and Sunakawa's (2010) work on what they call 'machine gaze' (rather than cameras). Expanding upon our brief remark from earlier about glancing in video games they suggest:

> The screen is necessarily a focus of gaze during game play, and the software has powerful gaze-shifting properties. Without moving their heads, eyes, or bodies, players can invoke—not only through symbols of speech but through the machine software—new gaze or sight parameters. They can rapidly switch between a first-person perspective and a third-person perspective, managing the actions of the self in relation to others in multiple ways. Gamers thus have individual and enhanced human-eye points of view on participation space to communicate and coordinate. (Keating and Sunakawa, 2010: 346)

As becomes apparent from their study, there is a form of glancing within the video games because players become adept at switching between the multiple camera perspectives that they are offered which then construct unusual virtual visibility arrangements of the game because it can be seen from multiple perspectival points. Switching between the multiple perspectives provided by game-cameras is distinct from the naked eye presented with a landscape, and also different again from the qualities of the physical screen on which game-camera perspectives are graphically rendered. Yet it is common enough to find settings where looking jumps between perspectives (e.g. looking in the rear view mirror of a car). Keating and Sunakawa's (2010) study also brings to light the importance of the 'participation cues' that happen in the space around the computer. In their study team players were sat together at a

LAN party and coordinated their moves by talk, gesture and handling of the keyboard (see also Sjöblom 2011). We can begin then to see how gameplay is produced from the skilled operation of cameras, the courses of actions that are generated by the game but also in the physical ecology of the computer display and the actions that surround it.

1.3. Interaction in virtual environments

Early studies of VEs, although not focussing on virtual camera perspectives per se, have examined how users look around, with and through these environments, focussing in particular on the relationship between newly emerging collaborative virtual environments (CVEs) and the transformation of pre-existing everyday visual practices in them (Hindmarsh et al., 1998; Hindmarsh et al., 2006). For instance, Hindmarsh and his collaborators (Hindmarsh & Heath, 2000; Hindmarsh et al., 1998) explored how CVE users establish (or more usually fail to establish) common points of reference through using the arms of avatars within the environment to point at objects.

Visual practices have been central from the outset to both the ways in which users of CVEs are embodied and how they engage in talk (Bowers, Pycock and O'Brien, 1996). Designers of CVEs themselves have supported these emerging visual practices by, for instance, developing interactional models to control visual and aural awareness (Benford and Fahlen, 1993). In early CVEs such as MASSIVE (Greenhalgh and Benford, 1995) and DIVE (Carlsson and Hagsand, 1993) the visual field was all the more pertinent because such environments were limited by simplistic graphical capabilities (e.g., simplistic geometries and low polygon count models, limited animation, basic light models and so on). The fragmentation and disruption caused by the primitive visual displays of rooms, corridors, furniture and so on, provided insights into ordinary visual practices, such as predicting the trajectories of action by others or establishing common points of reference (Hindmarsh & Heath, 2000).

Although multi-user games existed prior to the advent of CVEs and have been the focus of a number of studies (e.g., 'Multi-User Dungeons' (Dourish, 1998)), the introduction of ever more complex 3D VEs brings relevance to this prior work on visual practices in CVEs through turning toward learning and knowing the game (Sjöblom, 2008; Bennerstedt & Ivarsson, 2010). The development of online collaborative platforms in the rapidly evolving games sector has introduced new technologies that both advance the possibilities of playing together and the visual features within gameplay. Irani, Hayes & Dourish (2008) consider how users' self-perception and reciprocity of that perception are reshaped by the software construction of online collaborative VEs. They argue that rather than the ongoing attempt to simulate the properties and practices of 'real-world' environments, VEs might instead "create a novel platform for interaction" (Irani, Hayes, & Dourish, 2008: 195).

Moving beyond this literature, we see that game design for VEs, by its very ethos, has established all manner of unconventional and imaginative forms of visual practice. Situated on the border between a number of existing visual fields, game design and by extension gameplay itself, brings together the following topics:

> *New game-relevant visibility practices*. For example in a game a window allows a player to see an opponent inside a building and opponents will thus also try and avoid being seen through windows by change of trajectory etc.).

A combination of novel challenges for the player. Game designer innovation around pre-existing formats provide familiar resources for varied arrays of entertainment, skill, reasoning, narration, involvement, tactics, scoring, winning, losing, rules, aesthetics, nerves and more. The multiple concepts collected by the word 'game' always then require further definition.

Influence of cinema. The growing implementation by game designers of cinema's grammars for the display and recognition of characters, actions, events, scenes and more.

Cutting across all three of these topics are the uses of virtual cameras that both display the environment and create perspectives to play from, within and with. So, our study in this chapter enriches the existing body of research exploring visual practices in VEs by a direct examination of the overlooked relevance of virtual camera-work itself.

1.4. Cameras in games

The centrality of the camera to 3D video games is marked out in the very title of the ever-popular '*first-person* shooter' genre. In these games, the camera is predominantly positioned as if it were looking through the eyes of the player's avatar, which acts as a surrogate or proxy for their own view on the VE[1]. Nitsche has written the most comprehensive treatment of cameras in video games in his recent book "Video Game Spaces: Image, Play, and Structure in 3D Worlds" (2009). He begins by

[1] We note that while first- and third-person perspectives are common, second-person perspective games are extremely rare, the game Zato being one example in which the player sees their own embodiment 'through the eyes of' an enemy player (http://www.indiegames.com/2011/04/interviewing_indies_in_japan_s.html).

drawing upon research in film and media studies on the camera's role in cinema to chart how the positioning and movement of the camera is used to produce a sense of space and perspective in film. However, video games radically change the purposes of cinema camera-work when translated into a video game. Nitsche argues that video games hybridise two spaces: cinema's screen with its static audience and the built spaces of architecture with its mobile inhabitants. Effectively, computer games are built environments within which players can pan, tilt, dolly their cameras, swap cameras and swap lenses as part of their ongoing courses of action of exploring, chasing, dodging, killing and so on.

By contrast with cinema, where the choice of camera, its movement and its lens is tied to narration, emotion and other concerns of the film-makers, what the camera shows in video gameplay is restricted by its requirement to support the playing of the game (Nitsche, 2009). Consequently game players are not lead by the game-camera's movements like audiences are when they watch films, though interestingly gameplay has hybridised with film-making as Nitsche also notes (see also on spectators Taylor & Witkowski, 2010). The game player's handling of game-cameras is then closer to the physical camera-operator directing the camera than the spectator of a film. In addition, the field of possibilities offered by cameras in cinema and games is typically very different; while the cinema camera is tied to the physicality of the camera and the set (although this does not hold, say, for some special effects shots), the virtual camera as found in video games is not.

Having noted these critical differences, Nitsche takes us through the variety of modes in which cameras have been employed within video games since their inception. For 3D games, Nitsche identifies four dominant camera behaviours that are used within gameplay:

"Following camera." A third-person perspective that lies slightly behind the player looking over their shoulder or above their head.

"Overhead view." A different third-person perspective, top-down and similar to that from a plane or satellite.

"First-person point of view." Positioned as if the camera were on a headset in line with the player's eyes.

"Predefined viewing frames." The camera has pre-scripted behaviours at certain places within games (such as cuts or swapping to the view of your killer when you die (see Section 2.3)). (See Nitsche, 2009:102)

Following cameras position the virtual camera outside of the player's avatar such that the player can see their own avatar and something of the surroundings of the environment they are playing within. Of course, in many games the exact behaviour of third-person cameras may vary. In contrast, first-person cameras are tightly coupled with the player's manipulations of their avatar. At times these two may be mixed together, for instance in the role-play game *Fallout 3* (2008), a mixture of first-person and third-person camera positions are employed during its step-by-step combat system. Other games such as the *Resident Evil* series (1996-present) combine following cameras with static cameras set-ups in CCTV-like positions. The two games we are concerned with in this chapter fit within the first-person and third-person cameras respectively.

Virtual 'lenses' are further aspects of the properties of real-world cameras that are drawn upon in video games. At one level, this is used for aesthetic purposes by designers (e.g., lens flare giving the game a more cinematic feel). At another level, the properties of lenses are offered as instruments to players of first-person shooting games. As we will see later in this chapter players can select weapons whose telescopic lenses can then augment the player's capacity to investigate distant locations and, of course, aim accurately at remote opponents.

Finally, in various games the optical qualities of mechanical cameras are actively ignored to support novel gameplay. For instance, the first-person shooter *Aliens versus Predator 2* (2001) provides the player with a number of different alien avatars to select and play with that modify both the game's visual presentation and its play. Differing sizes of avatar offer the player different heights of game-camera perspective, movement abilities (e.g., an 'alien' can run across ceilings, whereas the 'human' cannot, thus corresponding to different camera capabilities). Avatar selection also alters the optical qualities of vision they provide (e.g., the 'predator' has limited colour vision thus modifying visual acuity). Breaking with conventions of camera height, movement and colour in these games helps remind us of the use of camera conventions as a resource in the intelligibility of play in the majority of 3D video games. In spite of this, we note that cameras are rarely topicalised in games. An exception is *Warco*[2] (2011), where the player takes control of a virtual video camera, records shots through its virtual viewfinder and then subsequently assembles them into news footage.

Having provided a brief survey of the variety of game camera possibilities we would now like to examine in detail how the camera is involved in the playing of two particular games.

---

[2] http://www.wired.com/gamelife/2011/09/warco

2. Study One—Playing in the first-person with a camera for looking and aiming: Counter-Strike

At the time of writing, Counter-Strike: Source ranked as one of the most popular games on Steam[3], a major games distribution platform. Originally developed in 1999 as a free modification to an existing first-person shooter called Half-Life, Counter Strike (CS) was subsequently released as a commercial game in its own right. Usually team-based, like many FPS games, CS involves navigating richly detailed 3D VEs with the goal of eliminating other players as well as achieving certain objectives (such as the 'capture' of a particular landmark within the game).

In order to orient the reader to typical aspects of CS that are engaged with by players, we offer an outline of the game experience as follows: A player selects a suitable server from a list of active game servers made available at the start of a session. The majority of play is performed online, with remote game servers acting as secure, authoritative hubs. The game itself is played on a particular set of 'maps,' each of which is effectively a self-contained 3D VE. Typically the player joins a server mid-game, 'dropping into' the action after they have selected a server (and therefore joining the current map and game type that is being played on that server). Players are then presented initially with a choice of two teams: terrorists and counter-terrorists. Players choose one of these 'sides' to play on when they enter, what is for them, the first round of the game. As CS works on a rounds-based system, the player

---

[3] http://www.steampowered.com

must wait until the current round has ended before they are ‘spawned’ with the rest of a team at particular points on the map.

As mentioned earlier, the characteristic perspective provided by FPS games is that of ‘seeing through the eyes of’ the avatar. The graphical rendering of this perspective is then presented on the physical display to the player, who may then attend to relevant parts of that display during play. In CS, like the majority of 3D games, interaction with the avatar is enacted via manipulations of keyboard and mouse. Thus, in CS, the perspectives provided by playing the game are, for the most part, constrained by the player-controlled movements of a humanoid-shaped avatar via the game’s keyboard and mouse interface.

The camera is sited on the avatar’s ‘head’ and directed by movements of the mouse. Players of CS typically use the keys ‘w,’ ‘a,’ ‘s,’ ‘d’ for forward, sidestep left, backpedal and sidestep right movements respectively. These actions are conducted concurrently with use of the mouse, producing streams of game-camera perspectives, rendered graphically to the screen. In this way the player may ‘peer into’ the VE through the physical display as ‘porthole’. It is through this ‘porthole’ that they observe and conduct game actions (running, jumping, shooting, etc.). Players’ planar movement in the VE is strongly tied to the orientation of the perspective created through this combination of mouse and keyboard—for instance, pressing a key for ‘forwards’ moves the avatar in the direction determined by where the centre point of the head perspective is currently direct, which itself is controlled by the mouse. Beyond this the movement of the game-camera is constrained by a number of aspects, such as the limitations of the avatar (e.g., it does not fly or pass through walls), and more generally, the game’s physics (e.g., the player can only travel at a ‘running’ pace), the geometry of the spaces travelled through, and the physical ability of the player to engage in dexterous manipulations of keyboard and mouse (Reeves, Brown & Laurier, 2009).

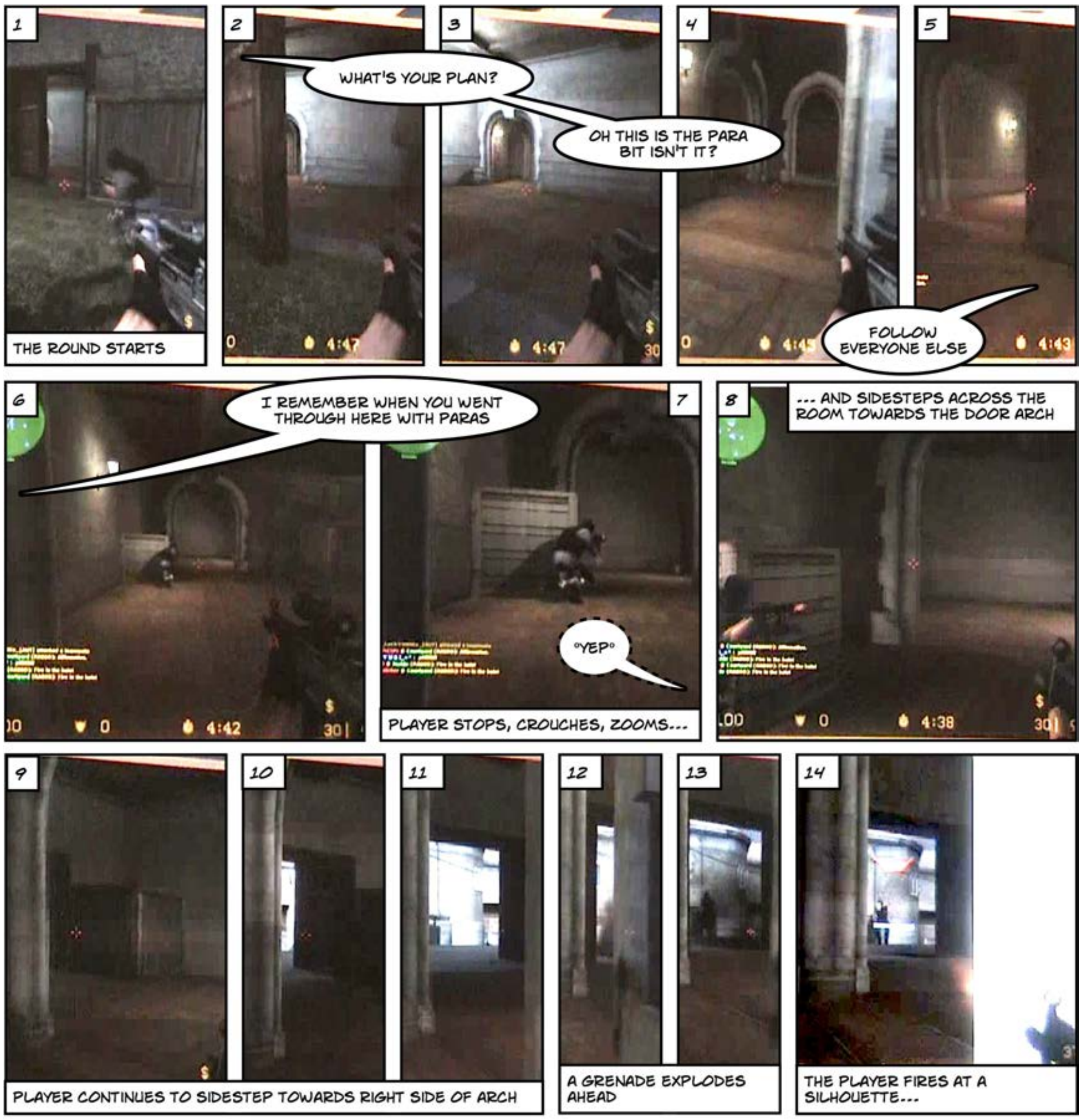
2.1. Establishing what is happening through moving the camera

Intrinsic to many FPS games is the strong coupling between a player's manipulations of their avatar and the game-camera perspectives that are produced through, and crafted in, those movements. The following sequence (transcript 1, panels 1-20[4]) introduces these aspects. The player has his brother sitting to one side (Reeves who is also video-recording the play). It is a common feature of gameplay, even where there are not dual controls, that there are peripheral participants in the game providing game-relevant (and game-irrelevant) talk (Lin & Sun, 2011; Aarsand & Aronsson, 2009). The round has just started, with the player having selected and joined a server, and waited for the current round that they have 'dropped into' to finish.

The player's camera, as they navigate a series of rooms is continuously and tightly coupled to the position and orientation of what would be their avatar's head, though, as we have noted already, this 'head' shot remains fixed straight ahead. With such tight coupling the player can precisely 'point' their camera around the environment during movement. Glossing this, we can say that throughout the series of rooms the player maintains camera perspective in a balance between looking for the action ahead and minimising their visibility to their opponents. Producing a perspective for looking at what is happening has to be weighed against pointing their weapon at potential threats.

To try and both look around but also have their weapon upon likely targets players become adept at ‘dollying’ the camera. However the film professional terms of ‘dolly’ and ‘camera’ need to be discarded in order to understand the players’ sensemaking within the game. To re-iterate, for the players there is the ongoing concern of both seeing what is happening in the game and avoiding being seen by their opponents in the game within the limits offered by the game’s graphics engine, their screen and, just as importantly by the game’s designed environment of exposed spaces, cover, vantage points, traps, etc.

As we have argued earlier, crucial to playing the game is arranging the camera’s perspective to allow for the ongoing surveying of the scene (Luff, Heath, & Jirotka 1999) in order to anticipate ‘what next.’ Turning to an episode of actual play, we see how the player approaches and then moves through a door, fires at the opposing side and then is killed:



Transcript 1: *Moving the camera through rooms and doorways. Player's speech bubbles are from the right and participant's from the left through CS transcripts.*

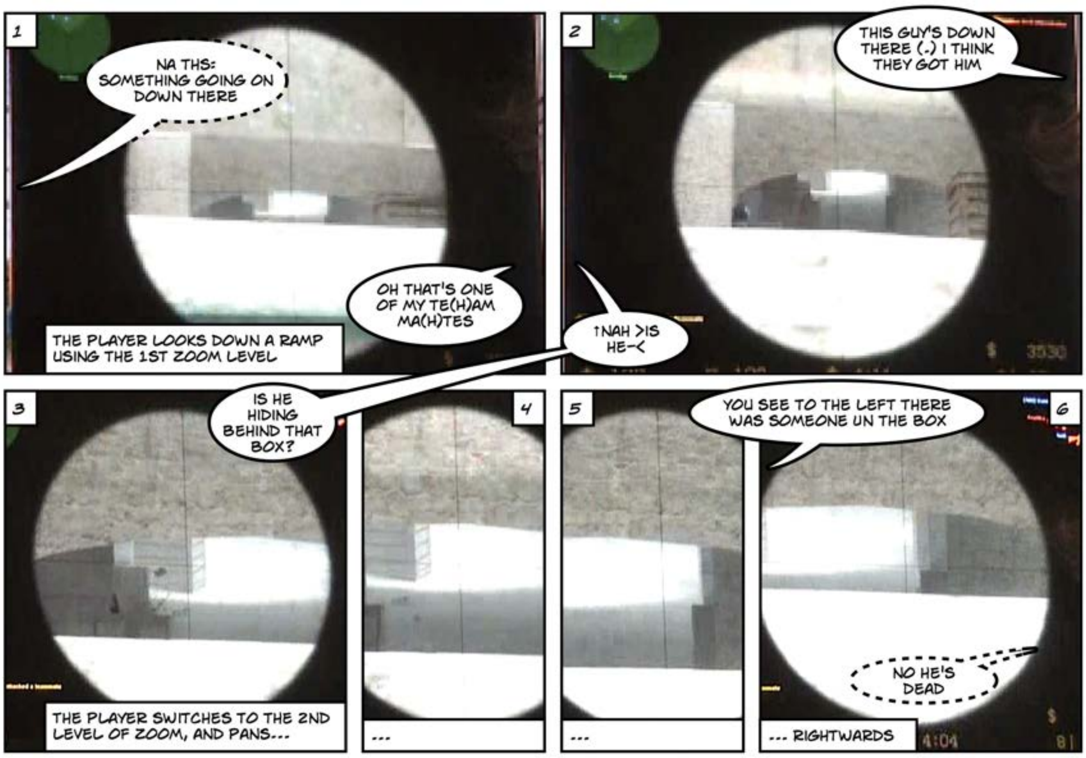
Rather than approach the doorway at right angles to it, the player and the remote player ahead have moved their avatars around to provide a view through the door that is ajar (panel 1-2). Through their adjustment of the camera perspective they can thus see something of the view into the next room from a distance and can move then onward along this line of sight. As their movement continues (panel 3) we see something of the payoff with the game-camera angle now offering at first a sliver, and then a full view of the next doorway. They very edging forward of play is produced in with the camera through carefully approaching openings that allow the player to both see ahead but also to be seen by distant opponents ahead.

Edging forward is done, not only with a concern toward visibility, but it is also built with a concern to bring the centre point of the camera to bear upon a features of the environment from where enemies are likely emerge from. In the fragment this is achieved by the player crouching down their avatar (panel 7), which thereby cranes the camera downwards while also increasing the accuracy of the weapon (shown by a narrowed target point). Near-instant firing upon the opponent when they become visible (and ideally before they see the player, and kill them) is a core component to this rapid round-based game. Thus the kind of camera-work being conducted expertly here by the player is combines the 'exploratory' (i.e., seeing where to go next) and the 'preparatory' (i.e., being ready to act).

The visual orientation of other players' avatars is visible to team members and is used as resources by the player. The team member's crouch ahead is recognised as a response to the next room, and, in turn our player performs a characteristic stop-crouch-zoom (panels 7-8) when they approach the same room. Having established a position that is surveying one perspective at this doorway, our player zooms from a slightly different angle thus increasing the overall territory ahead that is being surveyed by his team. Both their individual lines of sight are kept narrow enough to maintain minimal visibility for the players and thus making it harder for the opposing team to see them and shoot them. A wider field of scrutiny of the room ahead is thus co-produced and may then give them the advantage over their opponents.

2.2. Establishing what is happening through switching lenses

By selecting certain weapons with telescopic lenses, players are given further control over the virtual optics of the game. Each choice of perspective comes with a balance of properties that both augments and limits what can be seen. In transcript 1, panels 6-7, the player increases the accuracy of the weapon, and the fidelity of what is being looked at, but for all that the player gains they sacrifice other capabilities within the game. Looking through the telescopic sight reduces the player's field of view and the game automatically reduces the player's maximum speed of travel through the game spaces, modelling the limitations imposed by crouching while moving forward. While camera operators zoom as part of the practices of recording (Broth et al., this volume), the game-relevant practices of zooming via lens selection and game-camera perspective production by the player are meshed together—to use an in-game weapon like the sniper rifle is to use a lens selection. In transcript 2 we see the player manipulating their perspective through orienting the mouse in tandem with changing the optical qualities of the camera through zooming. Echoing the framing of telescopic lenses the sniper rifle is enclosed in a circle easily distinguishable from the regular rectangular of the screen. The sniper rifle has two levels of magnification that the player can use:



Transcript 2: *Zooming in and sweeping back and forth with the gun scope.*

In order to pass through to the most powerful level of zooming, the player must initiate a zoom to the first level (this is done simply using right mouse button clicks) and then cycle through: ‘normal’—‘zoom level 1’—‘zoom level 2’—‘normal’, etc. The optical functions of the weapon here are, of course, intended to support aiming and firing on targets that are some distance away. They allow more than this though, the different levels of magnification are used to conduct different kinds of visual inquiry that are relevant to producing distinct visibilities for different courses of action (Mondada, 2003). Cycling between the levels of magnification enables the player to briefly peer closely at a narrow territory of interest within a wider perspective.

Before the sequence in transcript 2, the player had been surveying the scene before then zooming in on the underpass. As transcript 2 begins the player zooms in and frames their shot so as to make relevant the tunnel and its opposite opening (panel

1) for their closer scrutiny. With their area of scrutiny aligned by the zoom, the onlooker provides the first analysis of what they can see: “ths: something going on down there.” There is then an, as yet unspecified, game action visible to him. The player responds with an oh-prefaced recognition (panel 1), his laughter through the recognition accounting for the fact that he had his cross-hairs tracking a figure below the bridge that he now recognises as one of his team-mates.

The player then makes further sense of what they can see “this guy’s down there I think they got him” (panel 2). ”[T]his guy” refers to one of the enemy that they had seen earlier below the bridge (which had led to the current inquiry into what was happening ahead). The onlooker responds, in disagreement with the conclusion that the enemy had been killed (panels 2-3). In the light of this disagreement over what has happened below the bridge the player swaps to the higher magnification zoom, which then provides a more detailed view than when the onlooker noticed that there was ”something going on.” The player’s use of magnification allows them both to see both movements and detritus. These are intelligible to the experienced player as the visible evidence of a close combat situation (e.g., when a player dies, their gun flies out of their hands (panel 3)). Because the greater magnification provides a restricted view of the area of interest, the player combines this level of zoom with a slow-paced sweep of the ramp and tunnel area (panels 3-5).

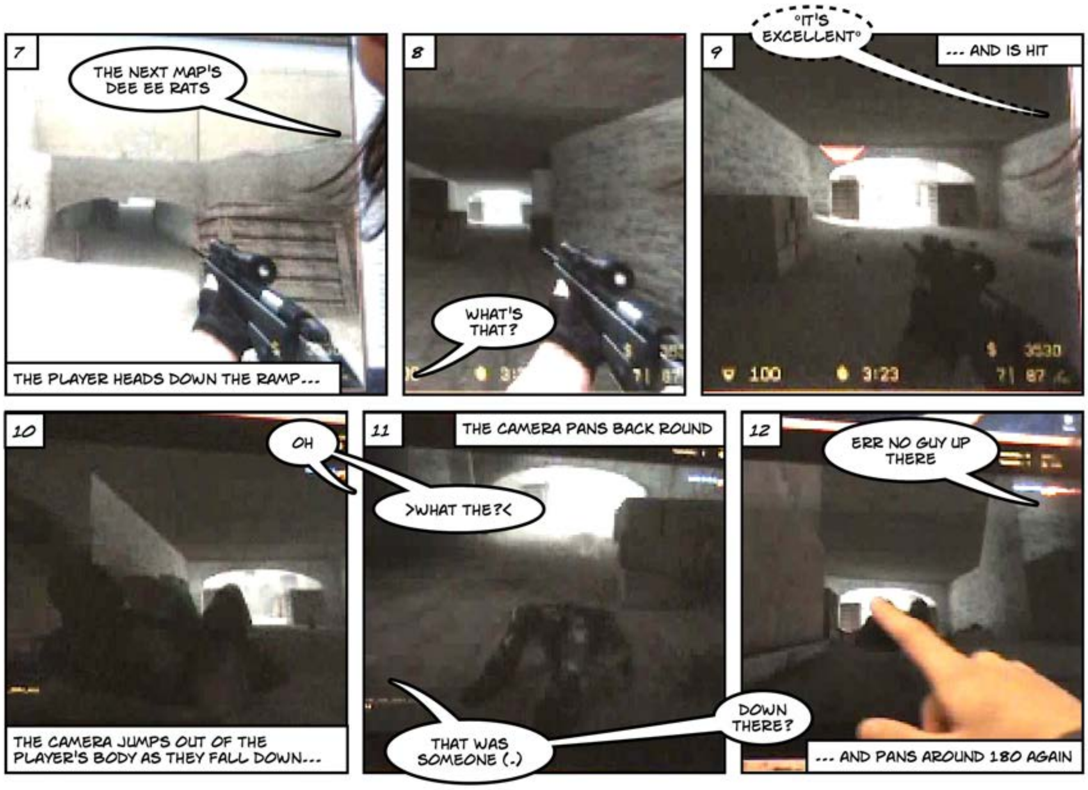
The sweep finished with the player lingering with his cross-hairs on the right side (panels 5-6) making that area relevant as the point where a new enemy player might yet emerge. In response to the right side being established as the player’s focus, the onlooker tried to redirect the scope to the left (panel 6). Even though they are trying to redirect the scope, the onlooker also accepts that that might no longer be the case by putting it into the past tense (“there was someone behind the box”). This ‘was’ provides some ambiguity over whether the enemy player ‘was alive’ or ‘was there but has now gone somewhere else.’ The player responds a correcting report on that opponent’s status (panel 6) supposing the onlooker has not understood what happened and ascribes the first understanding of an enemy player that was alive.

As we noted earlier, Macbeth (1999) tells us that “cameras can’t glance” (ibid:151) in the way that human eyes can but he also points out that we treat the camera’s limited field as having an ‘outside.’ In the player’s view of a video game there is also an ‘outside’ and the camera will be adjusted in response to projected everyday courses of action (e.g., an object moves out of shot going rightwards), audio (e.g., indicating enemy approaching), lighting effects etc. Irani, Hayes & Dourish (2008) also note this lack of peripheral vision in the “mediated looking” within VEs. The limited view of the game provided by the player’s single first person camera is supplemented in the game by further camera resources that are unlocked after the player is killed within the game and it is to those camera resources that we will now turn.

2.3. Cameras after death: what happened and what can others see

The penalty for ‘death’ within Counter-Strike is the player losing control of his avatar until the next round while the surviving players continue to fight. In correspondence with being killed the camera coupling is modified: the camera perspective ‘leaps out’ of the avatar and pans up to present an overhead, now third-person perspective of the player’s body. In terms of Nitsche’s (2009) typology these are “predefined viewing frames” (ibid:93): firstly, the camera is moved automatically to provide a brief dramatising view from above the dead body of the avatar. Secondly, the game offers the player, briefly, the perspective of the player that killed him. It thus supplements an already available Schützian ‘reciprocity of perspectives’ with a literal swapping of perspective, the instructional character of which is derived from the sequential organisation of its presentation to the player (i.e., first-person view of death, third-person view of the player's body and the view from their killer's perspective. In future rounds of the game the perspective of their opponent can then be used by players to provide tactical resources given the player (now) has a sense of what weapon the enemy team member was using and where they were shooting from. Once the predefined 'enemy player view' finishes, a new form control of the game-camera is returned to the player. The game-camera control offered to the dead is intriguing given our interests in the properties and uses of cameras within games. What is returned to the player is not control of any of the movement of the cameras; instead the player's relationship to the game shifts to something similar to that of an editor in the studio being able to swap between the cameras of his team members that remain playing. The metaphor ends there however because the player is not trying to broadcast or sequence these camera perspectives and usually flicks through them to pass the time while waiting for the next round to begin.

The automated camera provides an occasion for an analysis of what happened to lead to the player being killed. The next transcript is the continuation of transcript 2 where the player and onlooker were closely inspecting the underpass. When the player is killed we see the camera 'jumping' from the player's body and then providing a number of perspectives in rapid succession. The camera moves into position behind the player's body and looks down a tunnel in the orientation the player was facing when hit (panel 10), spins around to provide a view of where the player just came from (panel 11), and finally spins once again to trace a line of sight between the player lying dead on the ground and the enemy player (panel 12):



Transcript 3: *Finding out what happened through the automated camera.*

After the player is hit, he expresses his surprise: “oh >what the?<” (panel 11), a response cry that is also raising a question over what happened. Its question format showing that the source of the shots that killed him, is pending rather than being immediately reportable. The onlooker is the first to analyse the sequence of automatic camera perspectives automatically provided and says with a questioning intonation “that was someone (.) down there?” The directional sense of “down there” being provided in relation to the automated panning action of the camera. The onlooker’s candidate solution to who killed the player suggests the attack came from an as yet unseen enemy present in the tunnel itself.

However the camera algorithm was still unfolding and the alternate perspective of the camera offers the opportunity for the player to provide a subsequent solution to what happened as it shows a known sniping point: “err no guy up there”.

The player's "up there" located by his pointing (panel 12) to a location on the display where there is a balcony that overlooks the tunnel[5]. On the basis of these two examples we can grasp how the automated camera-work offer resources for resolving puzzles around 'what just happened?' and 'how did I get killed?'

3. Study Two—Playing in the third-person with a camera for looking around: Doctor Who

Before we deal with the camera set-up in this next game we need to also say a little bit about the workings of this game given it differs from CS. The Doctor Who Adventure Games (DW), belonging to the adventure game genre, are built upon the characters of 'The Doctor' and his then current companion 'Amy' from the popular and long-running BBC sci-fi series. In 2010 when the data was collected there were four games in the series. For each adventure the players arrive in a new place in a particular time period where one or another alien presence is endangering the locals. Quite what has happened and what the players need to do to complete the adventure is discovered in stages as the game progresses. Unlike CS they are not playing against other players online or locally, though in some of their challenges there are computer-generated monsters that can kill the Doctor or Amy. Also unlike the routinized familiarity of a small number of virtual environments for the CS player, a central task in adventure games (and DW) is navigation in an unfamiliar environment that, aside from the

---

[5] Note that this balcony is hard to see in the relatively low-resolution screen shots in this chapter.

villains and monsters, also contains a number of usually fatal hazards—this lends itself to our study in that it permits us to describe camera-work where the nature of the environment is uncertain (as opposed to CS where it is enemy player actions that produce uncertainty). During the game the player is usually moving the Doctor avatar with Amy following. For certain challenges, the avatar is swapped to Amy and / or the avatar works alone temporarily. If either of the avatars is killed then the game restarts from their last successfully completed section of the game.

The game can be played alone, or, with someone else involved. In the data that follows there are a father (Laurier) and his young son playing together. Their participation in the game is through a more varied set of participation pairs than the onlooker + player pair in the previous section. They shift between player + onlooker, player + planner, and player + player (Aarsand & Aronsson, 2009). It is worth noting that while siblings and peers playing together is common, play involving parents and children is relatively rare (Aarsand & Aronsson, 2009).

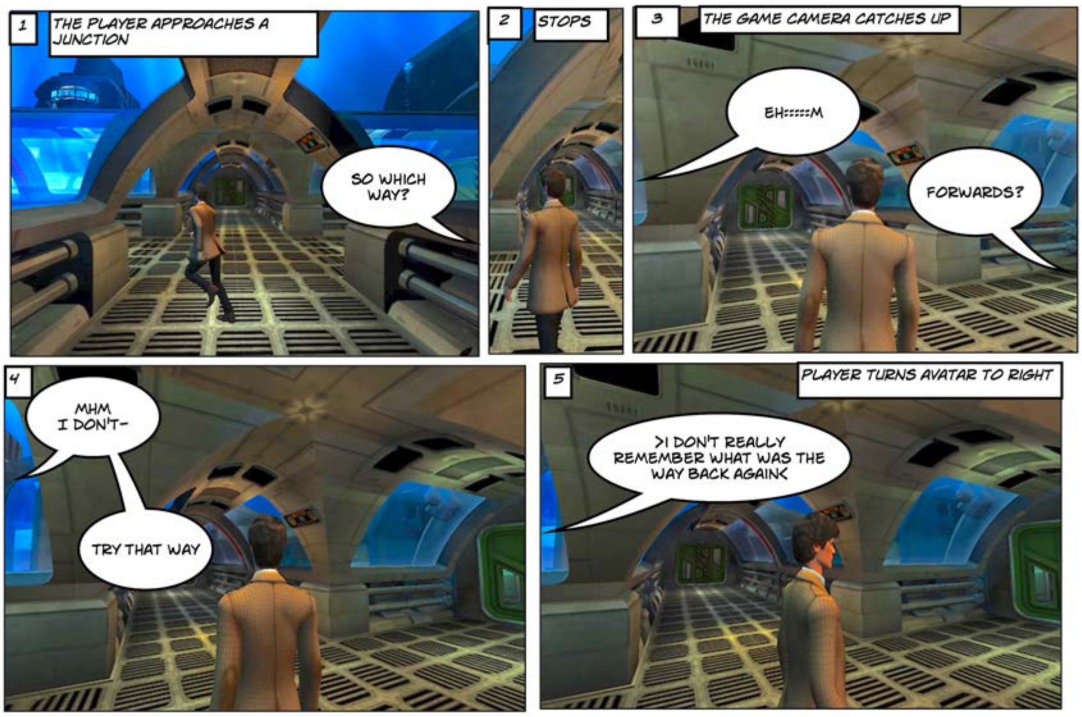
Like CS, players of DW use both keyboard and mouse manipulations to make moves within the game. In terms of the camera mechanics, the player's avatar is moved by one control (keyboard) and the camera is moved by the other control (usually the mouse). Characteristic of third-person perspective games, the camera's coupling to the avatar has inertial characteristics like a Steadicam and its distance from the avatar is also modelled on a 'spring' connection. When the character runs forward the spring stretches and when the character slows down the spring automatically pulls the camera closer to the avatar. This form of coupling also means that the camera's position is such that it floats upwards when following the player as they run through the environment. However, in a significant departure from the first-person perspective, the camera can be swung around by the player, out of synch with the body movements of the avatar. It can be panned, tilted and craned around the avatars allowing views from the ceiling down to the floors and vice versa, producing views of the characters from the front, sides and rear and so on. Such manipulations of the game-camera by the player are, at points, major resources in progressing in the adventure.

While the game-camera's operation is not cinematic it does require the players to organise two kinds of actions: firstly, camera-actions (e.g., panning, tilting and rolling) that produce appropriate exploratory or preparatory perspectives for looking around, or for next actions, and secondly, the movement of the avatar (and thus the camera) through the VE. These are controlled from separate computer inputs (keyboard and mouse) and the player and co-player share or swap the controls from time-to-time according to the ongoing sequence of action.

3.1. Convergent and divergent looking

We join the players in this game where we left off with CS in the midst of a collaborative examination of what is displayed on screen. While much of the time they share a focus in the gameplay, they do not always converge on the same thing. In fact the possibilities for divergent orientations toward what is being displayed on the screen is all the more apparent in them playing DW because the father is more closely involved than was the case of the player and his brother in CS. What the father takes on for much of the time is the navigational work of establishing where in the complex labyrinth of the game they are and where they need to go to next. In the meantime the son is devoted to controlling the avatar and from time-to-time operating the game-camera.

We join them as they are in the midst of trying to find their way around a maze of tubular corridors in a seafloor station of some kind. The tunnels have windows allowing them to see the wider undersea neighbourhood of buildings and connecting corridors.



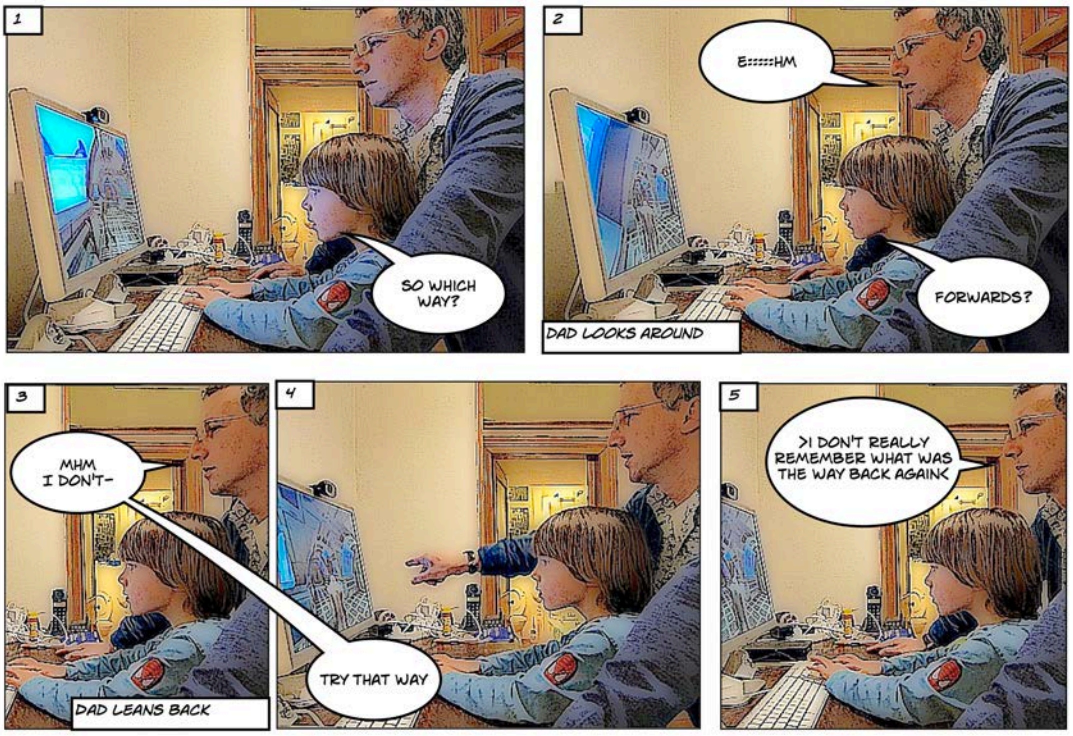
Transcript 4a: *The following camera & turning the avatar. Player's speech is from the right and co-player's is from the left throughout the transcripts of DW.*

When the camera is simply following the avatar, it provides a straight-ahead centred perspective. The avatar's arrival at the junction makes relevant an inquiry and the player asks "so which way" (panel 1) thus also directly requesting help from the father as a co-player. The directional help being requested makes relevant the paired category in the adventure game of player + navigator.

We also see two behaviours of the camera's automated coupling to the avatar's movements. The camera is initially stretched behind the Doctor as he runs and then, when he stops at the junction, it slows down to catch up with him (panels 1-

3). Secondly, when the player turns the doctor to the right the inertia delays the camera from swinging immediately behind him (panels 4-5).

While transcript 4a gives us a sense of the camera set-up within the game, if we swap to a view of the players around the screen we can see the more of the physical-virtual ecology of the father, son and game-camera[6]:



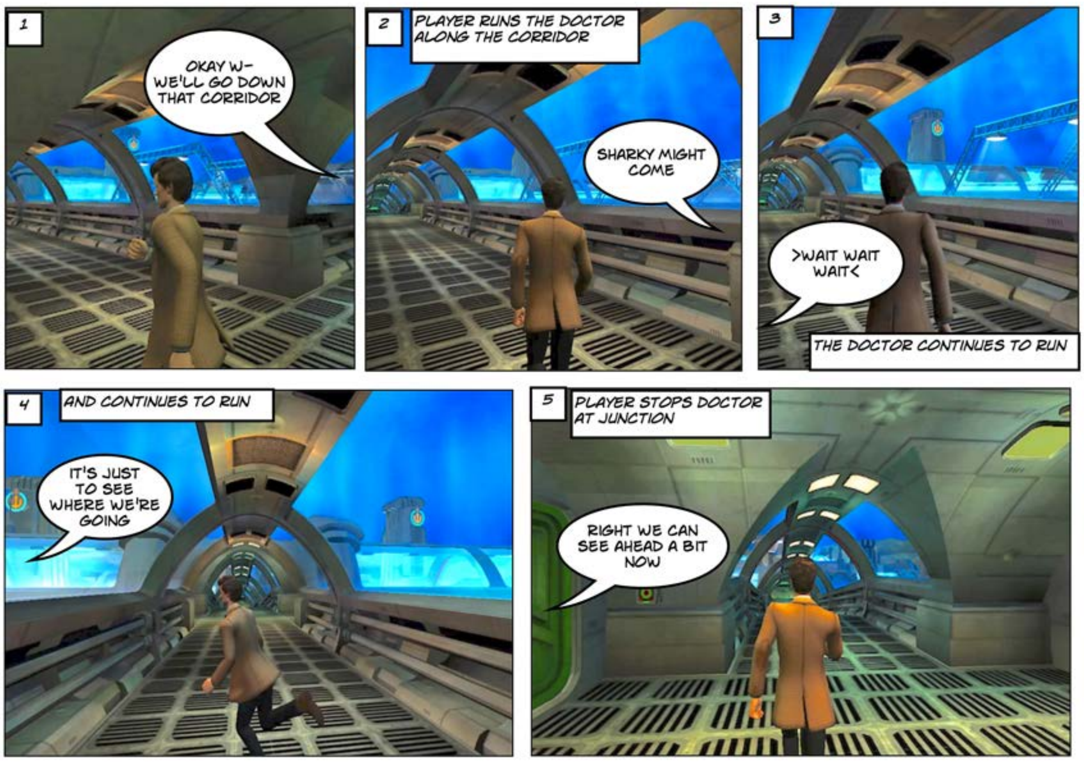
Transcript 4b: *Pointing over a screen to indicate direction.*

The possibility of gesturing around and over a shared screen is common across a wide variety of gameplay (Keating & Sunakawa, 2010) and video screens more widely (Heath & Luff, 2000; Hindmarsh & Heath, 2000). The father and son are in a huddle around the computer with the son controlling the keyboard and mouse, and the dad close behind him, looking over the top of his head. As player and co-player they

---

[6] This is something we have only briefly see in CS, for instance transcript 3, panel 12.

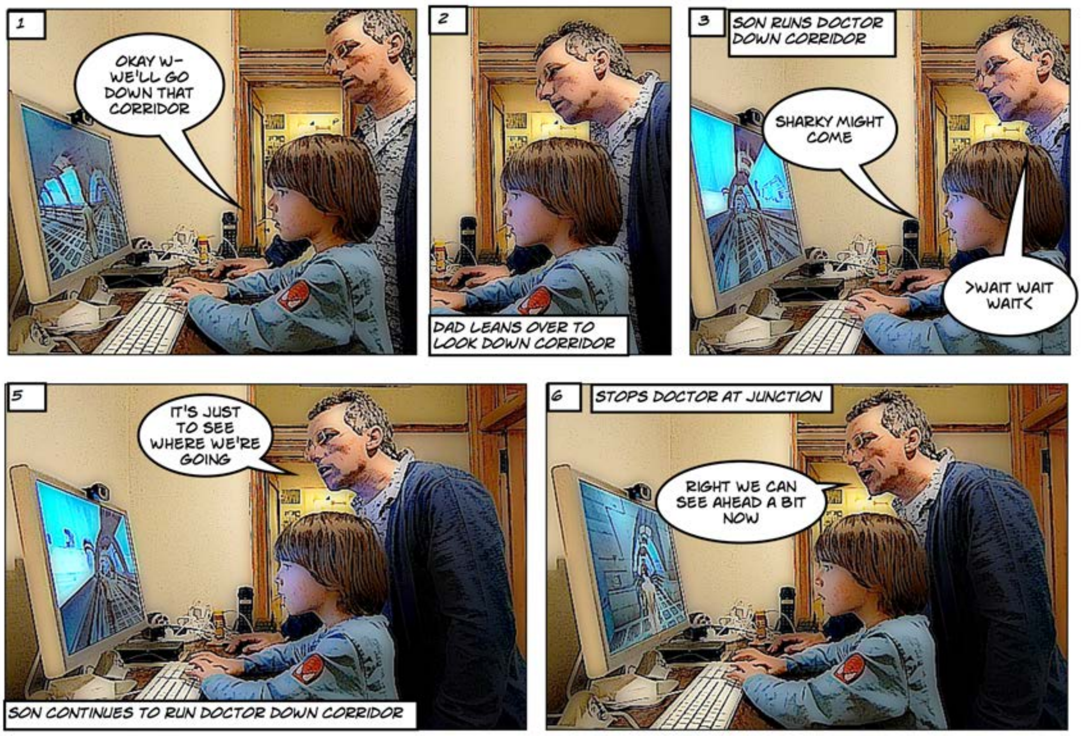
can monitor the screen for what is currently happening. What the co-player can do, given their hands are free from the controls is then also gesture across, at, on and around various features that emerge in the game's environment by virtue of their graphical rendering on-screen. We thus see the co-player suggesting "try that way" (transcript 4b, panel 4) while simultaneously pointing toward the right hand corridor. In this fragment of the gameplay the camera is 'transparent' (i.e., it is not topicalised); the co-player is giving a direction gesture not all that different from one that we might see from a passenger in a car toward what can be seen of the junction ahead through the windscreen. As mentioned earlier, the screen provides a shared 'porthole' into the VE that is ongoingly aligned for the two players. Because the players are already accessing the game through a 'porthole', much of the time they do not need to establish an alignment of perspectives in the way that pedestrians exploring urban environments do. The latter may be facing in different directions or be differently placed in relation to lines of sight and so on that then require work to align their visual orientation toward the environment (see Mondada, 2008 and Laurier & Brown, 2008). Having said that the 'porthole' or 'windscreen' produces a shared perspective, what we began to see in transcript 4b was the possibility of non-aligned / divergent orientation toward the features of the environment. This becomes clearer in the second data fragment:



Transcript 5a: *Running through the tunnel.*

In transcript 5a, panel 1 the player accepts an earlier proposal from the co-player to select a particular corridor. He warns of one of the threats they face when they travel in the windowed corridors (e.g., “sharky” a shark that smashes the glass and drowns them). The player’s concern is around being visible to the shark and so he hurries through the glass corridor. The co-player, in the face of the avatar rushing through the glass corridor, requests that he wait (panel 3) and then elaborates on why he is instructing the player to wait as navigation-relevant (panel 4). The player continues thereby, ignoring the request to wait and then ignoring the account for the request, continues running his avatar still further down the corridor (panel 4). When the player does finally halt it is at a junction out of sight of “sharky” because there are no immediate windows. The co-player is trying to do reconnaissance work that is being denied by the player whose course of action is to leave behind the windows as fast as possible to avoid the shark attack.

If we shift toward how this plays out within the game's ecology of screen and controls, we can see that the co-player is also looking at the peripheral area of the screen (transcript 5b, panels 2-5) where the windows are:



Transcript 5b: *Looking at different areas of the screen.*

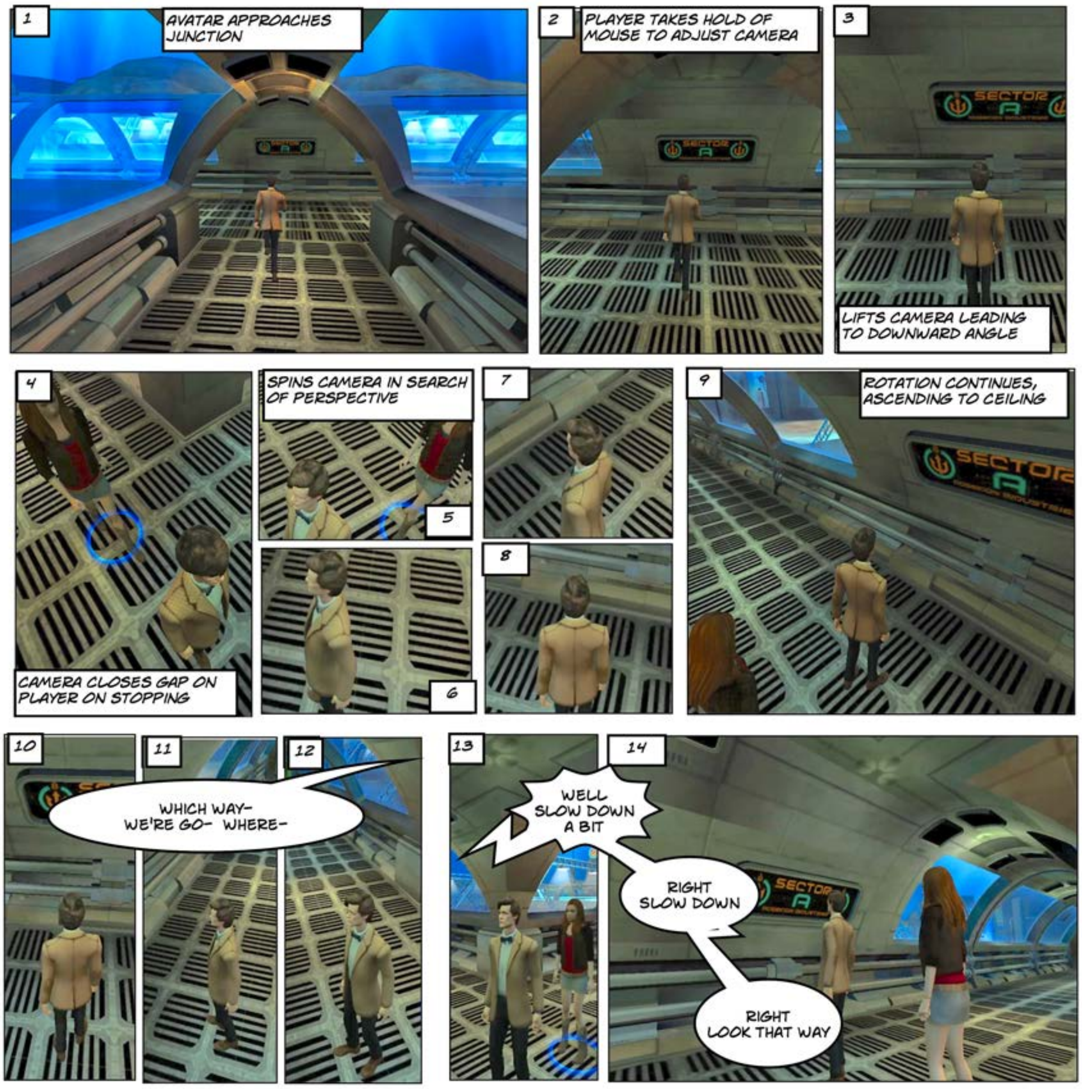
As we noted above the lack of alignment is generated by the player's desire to avoid "sharky" and thus also by having to 'drive' the avatar and is thus attuned to the corridor ahead for upcoming junctions. The corridor ahead is kept centre screen by the following game camera. The co-player, in the meantime, is inspecting the sides of the screen, examining the buildings and other corridors visible through the windows of the corridor in the hope of working out where they are in the maze of corridors.

In Chuck & Candy Goodwin's work (Goodwin, 1997; Goodwin & Goodwin, 1996) on the distributed work of controlling airports, staff located in the operations rooms distribute the work of monitoring the airport across multiple screens showing multiple CCTV views. At certain critical junctures they can draw upon their multiple cameras to find out what is causing a particular problem in the airport. For the game players here, while they can distribute the work, there is only one screen offering one rendered graphical game-camera perspective. However what the player and co-player can do here is to scrutinise different regions of the screen.

3.2. Looking around with the camera

So far we have examined elements of how the automatic following-camera perspective of the game is used by the players, and how the distribution of the visual tasks is then accomplished around the screen, what we will turn to now is the direct manipulation of the camera using the mouse. This sits in useful contrast with the first person shooter game (CS), where we saw how camera-work was effectively one and the same as avatar actions. The differences in DW allow us consider what happens when the camera movement can be decoupled from the avatar movement.

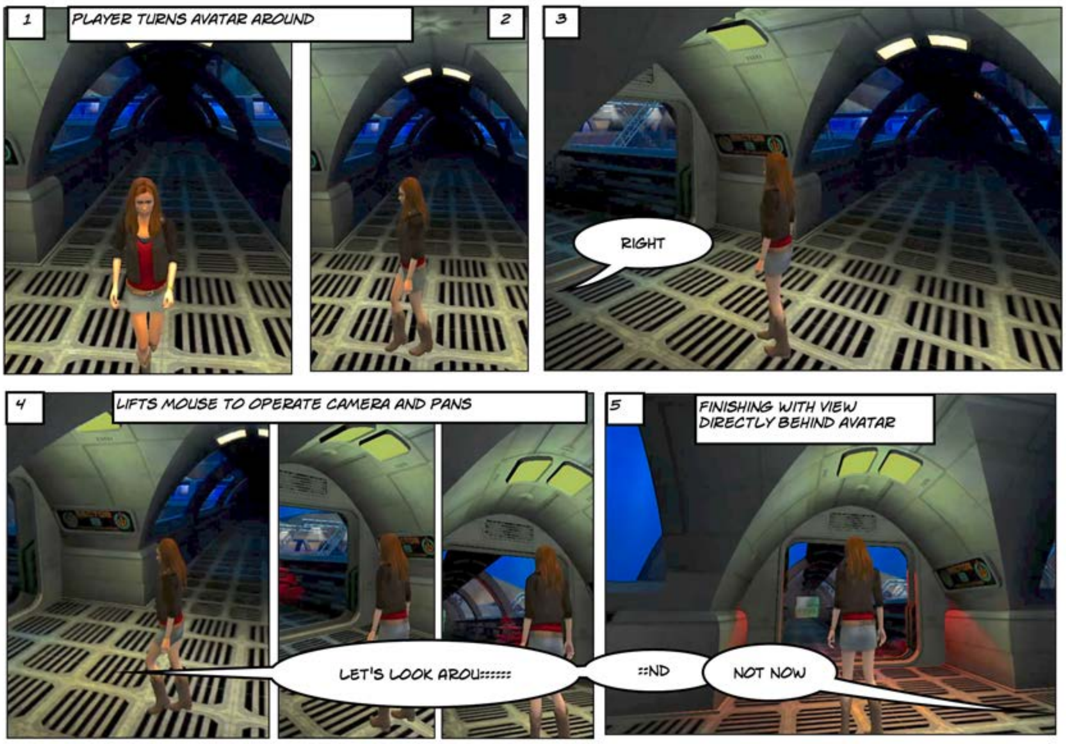
In transcript 7 the player is approaching a T-junction at the end of a corridor and, this time, rather than moving his avatar, he switches to moving the camera to look around from a fixed point in the midst of the junction:



Transcript 7: *Fast panning the camera.*

When the player engages the mouse the camera is lifted toward the ceiling (panels 2-3) producing a view from above. Immediately and rapidly after this, the player begins a somewhat inelegant sweeping scan of the environment at the junction by panning, tilting and craning the camera with the planar movements of the mouse (panels 5-7). The player continues the camera's rapid movement whilst formulating a halting question about direction (panels 10-12). As the co-player interrupts with an emphatic solution to the player's troubles, the player himself is already repairing the pan of the camera by lowering it back down and slowing its movement.

As it was with the first person shooter, handling the camera is not only for exploring the VE, it is also for preparing and positioning the camera perspective to provide an adequate view for future action. This becomes obvious in sections of the game where, similar to CS, the players are struggling against an opponent that can kill them (rather than navigating or solving puzzles). In transcript 8 they have to advance through the corridor while remaining in a moving red spotlight because the shadows are fatal. The speed of the spotlight has been designed as a challenge because it does not move at quite the same speed as the avatar and so requires adjustments of pace. The son finds this too difficult and asks his father to take over the controls of the game. They have discovered (having been killed three times in prior attempts) that the traverse requires precise timing in entering into the moving spotlight. Timing their actions to the spotlight is made easier by adjusting the camera perspective to look down the corridor toward the moving spotlight. As we join them, the corridor ahead is not the one they must take and the avatar has just been turned away from it:



Transcript 8: *Panning camera to prepare for next action.*

When the avatar is turned on the spot to face the tunnel with the spotlight (panels 1-3), the automatic following-camera leaves the players with a perspective from the side of the avatar. This difficult corner movement is part of the design of the game that requires bringing together avatar and camera controls in a more accomplished coordinated move. In a slow and steady movement the player pans the camera to provide a view from behind the avatar. The camera pan needs to be even more carefully crafted in that it also requires craning downward for a long view down the corridor with the moving spotlight (panels 4-5). The player accentuates the slow speed in his accompanying account of his camera movement by extending his utterance “aro::und” along the course of this movement. There is perhaps a hint of instruction in this given his son’s repeated rapid pans (as witnessed in the previous section). On completing both his utterance and the movement, the camera is brought to rest by the player in position behind the avatar. Located behind and relatively low, this ‘preparatory’ camera-work does then provide the longer view of the moving spotlight. Thus, it is then perfectly positioned to prepare them for the next action of moving the avatar forward into the spotlight. The co-player, tracing and projecting the trajectory of the curve of the camera, and the movement of the spotlight, warns the player not to make his move yet (panel 5).

Within this second study we have elaborated and expanded on the points of the first study, examining: how the player and co-player undertake distributed gameplaying tasks on a shared single screen; how the camera itself becomes topicalised within the game; and how the camera is operated to prepare for next actions within the game.

4. Summary and closing points

While this chapter has explored fragments of play from only two of the myriad video games that are currently played around the world, it has begun to uncover camera-work that is common across many of these games. The games that we have considered here are ones where the camera's view is also used to analyse the visibility of the player to their opponents; our study has revealed how game-cameras mediate the visibility of a virtual environment and what is happening in that environment. Both algorithmic and human action are involved in the production of these camera perspectives and both provide pertinent resources for play. In bringing this chapter to a close we summarise key analytic points raised by the two studies, and their implications.

In our first study we examined three features of camera activity in Counter-Strike (CS). We began by exploring how perspectival control of the camera is a skill that the player must master: i.e., successful play of the game often relies on successful camera-work. Secondly, we saw how important the role of non-first-person (often quite 'cinematic') camera perspectives are for players in terms of resolving normal game troubles such as 'how did I get killed?' Finally, we explored how zooming provided optical modifications of regular camera perspectives, offering new ways of scrutinising the VE, and, relatedly, the ways in which these optical characteristics, in affording different visual acuities, were enmeshed in the routine forms of 'looking' and 'glancing' enacted by the player. Zooming, in particular, may be seen as developing a form of camera 'glancing' that is very unlike un-augmented human vision. (We do not routinely zoom in on distant details, though we could perfectly well carry binoculars with us everywhere.)

In our second study of the Doctor Who Adventures Game (DW) we saw how the following-camera with separate controls is brought to bear in a range of practices. Playing the game together required an ongoing division of labour between player and co-player in different elements of the gameplay (e.g., moving the avatar, panning the camera, scrutinising the screen). In exploring the movement of the camera and how it connected to the movement of the avatar, we highlighted convergent and divergent visual inquiries of the game as visible on the screen. Throughout both studies we examined the manipulation of the camera to find and produce perspectives that are relevant to the 'next move,' in order to be used in navigating through the labyrinthine environments of both games.

The main contribution of this chapter is in explicating *gameplaying in and as* a matter of 'camera-work.' At the same time our chapter contributes to the growing set of studies of seeing with a camera (see the introduction to this volume) and is of a particular form where the phenomena seen with the camera can only be seen via a game-camera. Without the virtual camera there is no playing of the game of Counter-Strike, nor Doctor Who. It is thus a particular kind of camera whose centrality means that it is easy to overlook the fact that it is a camera at all. Indeed, the moment-by-moment, practical work of operating the game-camera has remained largely ignored as a foundational feature of play within video games studies. From our two case studies in this chapter we have begun to describe how the panning, dollying and zooming of cameras are sequentially organised in, and draw upon the sequential organisation of, the playing of video games. If we wish to understand video practices within video games, then we need to move beyond listing their many and varied configurations (e.g., Nische, 2009) to explore how these technical configurations are drawn upon, assembled and reconfigured, in and through gameplay.

The other side of this contribution is in a deeper understanding of *embodiment* in virtual environments. While there is a significant body of literature exploring the character of embodiment in VEs (e.g., Hindmarsh et al., 1998), typically such work neglects or glosses the role of the virtual camera; for instance within Hindmarsh, Heath & Fraser (2006) the 'view' or 'viewpoint' is presented yet not unpacked as a resource or topic for the user.

Finally, in unpacking gameplaying practices in and as virtual camera-work, and deepening understandings of embodiment in VEs, this chapter also helps us expand and reconfigure prior understandings (e.g., Irani, Hayes & Dourish, 2008) about how 'looking' is practically done around virtual environments. The idea of seeing-with-the-camera has become ever more relevant and interesting for game players, and the designers of games (Nitsche, 2009) and is perhaps at odds with the initial idea of the simulation of seeing-with-human-eyes that caught the attention of earlier researchers on VR environments (Hindmarsh, Heath, & Fraser, 2006).

Acknowledgements

Our patient family members: Reuben and Philip. Part of this research was funded by the EPSRC ref. EP/F03038X/1. For allowing us to use their visual images from their games: Steam (http://www.valvesoftware.com/company/index.html) and BBC Doctor Who Adventure Game (www.bbc.co.uk/doctorwho).

References

Aarsand, P. A., & Aronsson, K. (2009). Response cries and other gaming moves—Building intersubjectivity in gaming, *Journal of Pragmatics*, *41*(8), 1557–1575.

Benford, S., & Fahlén, L. (1993). A spatial model of interaction in large virtual environments. *Proceedings of the Third European Conference on Computer-Supported Cooperative Work* (pp. 109–124).

Bennerstedt, U., & Ivarsson, J. (2010). Knowing the Way. Managing Epistemic Topologies in Virtual Game Worlds. *Computer Supported Cooperative Work (CSCW)*, *19*(2), 201–230.

Bowers, J., Pycock, J., & O'Brien, J. (1996). Talk and embodiment in collaborative virtual environments, *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 58–65).

Broth, M. (2004). The production of a live TV-interview through mediated interaction. In C. van Dijkum, J. Blasius, H. Kleijer & B. van Hilten (Eds),. *Proceedings of the Sixth International Conference on Logic and Methodology.*

Broth, M. (2009). Seeing through screens, hearing through speakers: Managing distant studio space in television control room interaction*, Journal of Pragmatics*, *41*, 1998-2016.

Carlsson, C., & Hagsand, O. (1993). DIVE A multi-user virtual reality system. *Virtual Reality Annual International Symposium, IEEE*. (pp. 394–400).

Dourish, P. (1998). Introduction: The state of play. *Computer Supported Cooperative Work (CSCW)*, *7*(1), 1–7.

Goodwin, C. (1997). Transparent vision. In E. Ochs, E. A. Schegloff & S. A. Thompson (Eds.), *Interaction and Grammar* (pp. 370-404). Cambridge: Cambridge University Press.

Goodwin, C., & Goodwin, M. H. (1996). Seeing as situated activity: Formulating planes. In Y. Engeström & D. Middleton (Eds.), *Cognition and communication at work* (pp. 61-95). Cambridge: Cambridge University Press.

Greenhalgh, C., & Benford, S. (1995). MASSIVE: a collaborative virtual environment for teleconferencing. *ACM Transactions on Computer-Human Interaction*, 2(3), 239–261.

Heath, C., & Luff, P. (2000). *Technology in action*. Cambridge: Cambridge University Press.

Hindmarsh, J., Heath, C., & Fraser, M. (2006). (Im)materiality, virtual reality and interaction: grounding the “virtual” in studies of technology in action. *Sociological Review*, *54*(4), 795–817.

Hindmarsh, J., & Heath, C. (2000). Embodied reference: A study of deixis in workplace interaction. *Journal of Pragmatics*, *32*, 1855-1878.

Hindmarsh, J., Fraser, M., Heath, C., Benford, S., & Greenhalgh, C. (1998). Fragmented interaction: establishing mutual orientation in virtual environments. In *Proceedings of the 1998 ACM conference on Computer supported cooperative work* (pp. 217-226). ACM.

Irani, L. C., Hayes, G. R., & Dourish, P. (2008). Situated practices of looking: visual practice in an online world. *Proceedings of the 2008 ACM conference on Computer supported cooperative work* (pp. 187-196).

Lin, H., & Sun, C.-T. (2011). The Role of Onlookers in Arcade Gaming: Frame Analysis of Public Behaviours. *Convergence: The International Journal of Research into New Media Technologies*, *17*(2), 125–137.

Keating, E., & Sunakawa, C. (2010). Participation cues: Coordinating activity and collaboration in complex online gaming worlds. *Language in Society*, *39*(3), 331–356.

Laurier, E., & Brown, B. (2008). Rotating maps and readers: praxiological aspects of alignment and orientation*. Transactions of the Institute of British Geographers*, *33*(2), 201–216.

Luff, P., Heath, C., & Jirotka, M. (1999). Surveying the Scene: Technologies for Everyday Awareness and Monitoring in Control Rooms. *Interacting with Computers*, *13*(2), 193-228.

Macbeth, D. (1999). Glances, trances and their relevance for a visual sociology. In P. L. Jalbert (Ed.), *Media Studies. Ethnomethodological Approaches (*pp. 135-170). Lanham MD: University Press of America.

Mondada, L. (2003). Working with video: how surgeons produce video records of their actions. *Visual Studies*, *18*(1), 58-73.

Mondada, L. (2009). Emergent focused interactions in public places: A systematic analysis of the multimodal achievement of a common interactional space. *Journal of Pragmatics,* 41(10), 1977-1997.

Nitsche, M. (2009). *Video Game Spaces: Image, Play, and Structure in 3D Worlds*. London: MIT Press.

Reeves, S., Brown, B., & Laurier, E. (2009). Experts at play: understanding skilled expertise. *Games and Culture, 4*, 205-227.

Sjöblom, B. (2011). *Gaming interaction: Conversations and competencies in internet cafés*. Linköping University, Linköping.

Sjöblom, B. (2008). Gaming as a Situated Collaborative Practice. *Human IT*, *9*(3), 128–165.

Sudnow, D. (1972). Temporal Parameters of Interpersonal Observation. In D. Sudnow (Ed.), *Studies in Social Interaction* (pp. 259-279). New York: Free Press.

Taylor, T. L., & Witkowski, E. (2010). This is how we play it: what a mega-LAN can teach us about games. *Proceedings of the Fifth International Conference on the Foundations of Digital Games* (pp. 195-202).