

Why Can I See My Avatar? Embodied Visual Engagement in the Third-Person Video Game

Daniel Black¹

Abstract

This essay seeks to answer two questions raised by the success of video games where the player looks at the character she is playing rather than seeming to inhabit the same coordinates as the character within the game space. First, why is the experience of playing these games not innately inferior to that of playing games with a first-person point of view, given that the sense of being a character sensing and acting inside the game space could be expected to be much stronger when the character's body seems to be one's own rather than a separate entity in the game space? And second, if the first-person point of view is so "immersive" and provides such a sense of being "inside" the representational space as is sometimes claimed, why has it never been so prominent in other audiovisual entertainment media such as film and television?

Keywords

perspective, point of view, embodiment, immersion, first-person, third-person, representation, aesthetics

¹ School of Media, Film & Journalism, Monash University, Caulfield, Victoria, Australia

Corresponding Author:

Daniel Black, School of Media, Film & Journalism, Monash University, Caulfield, Victoria 3145, Australia.
Email: daniel.black@monash.edu

Introduction

Existing theorizations of the video game-playing experience have, on various grounds and in various terms, been criticized for focusing on the visual aspects of games and so not paying sufficient attention to the fact that video games are not simply about looking.¹ While such complaints are justified, it is important to understand that the problem with older, ocularcentric accounts of gaming is not the fact that they focus on the visual in itself; as reflected in the very term “video game,” the visual is of critical importance to the game-playing experience. The problem is that they tend to perpetuate a set of misconceptions about visual perception that see it as disengaged, immaterial, and distinct and isolatable from physical action and other kinds of sensory experience. On the contrary, video games are dependent for their operation on the fact that we always have an embodied, multimodal, and active engagement with what we see, which can—to at least some degree—cause us to engage with a two-dimensional (2-D) visual representation as if it were a real, physical space of action.

In order to explore how such an engagement is possible, in this essay I will first highlight the degree to which the representational regime of the three-dimensional (3-D) video game is dependent upon the interaction of a number of discrete and importantly different viewpoints. I will then go on to consider the ways in which the body of the game player can function as a site at which different experiences of vision and action can be combined into a relatively seamless whole, producing as a result a sense of distributed embodiment and a distributed capacity for action that crosses from physical into simulated space.

Key to understanding this effect is the avatar, which acts as the player’s proxy in the game world: a point of articulation between the player and a simulated environment with which she cannot interact directly. When the player looks at the avatar, to what extent is it possible for this visual relationship to create a sense that she is sharing a game character’s experiences and actions? Does the player experience a seamless melding with the avatar, such that she feels as if she is in the virtual environment and acting upon it directly? Or is the avatar in reality experienced as simply a tool employed by the player to solve problems represented on a screen? Both of these characterizations seem too simplistic to be entirely true, but a more nuanced account requires that the game player simultaneously have quite different, even irreconcilable, embodied experiences.

Being in the Game

The first-person shooter (FPS) might be considered the exemplary “immersive” video game genre, producing an experience of “being” the protagonist of the game, seemingly situated at the same coordinates in the game space as the game character’s body and seeing and hearing the game world through the character’s eyes and ears.² At the same time, however, there have been plenty of games that do not present the

player with a game character's first-person point of view. Some of these games do not seek to create a realistically simulated 3-D space at all, but many do. For example, some of the most successful video game franchises, such as *Grand Theft Auto* (1997), *Assassin's Creed* (2007), and *Tomb Raider* (1996), and most of the massively multiplayer online role-playing games such as *World of Warcraft* (2004) seek to create (and succeed in creating if sales figures are anything to go by) a sense of acting upon a 3-D game space, and yet they primarily utilize a third-person viewpoint, in which the player viewpoint is situated only *near* the game protagonist, looking at the avatar as a separate entity in that space. When considering the possibility that a given regime of visual representation might allow a game player to feel that she in some sense is interacting with a simulated environment while playing, two questions are raised by the success of games where the player looks at the character she is playing rather than seeming to inhabit the same coordinates in the game space. First, why is the experience of playing third-person games not innately inferior to that of playing first-person games, given that the sense of being a character sensing and acting inside the game space could be expected to be much stronger when the character's body seems to be one's own rather than seeming to be a separate entity in the game space? And second, if the first-person point of view heightens realism by providing such a sense of being "inside" the representational space, why has it never been so prominent in other audiovisual entertainment media such as film and television, which are just as technically capable of utilizing it as video games?

James Newman has influentially argued that a tendency to focus on issues of representation in video games is misguided (Newman, 2002); an unhelpful inheritance from the analysis of other media such as film, a focus on the visual, and in particular how it might work to create identification between player and game character, obscures the specificity of the video game experience:

I want to suggest that, for the controlling player during gameplay sequences, the notion of "character" is inappropriate. Here, the "character" is better considered as a suite of characteristics or equipment utilized and embodied by the controlling player. The primary-player-character relationship is one of vehicular embodiment. In suggesting this model, I seek to challenge the notion of identification and empathy in the primary-player-character relationship and, consequently, the privileging of the visual and of representation-oriented approaches. (Newman, 2002, ¶ 3)

This argument suggests that, rather than identifying with the character in a way reminiscent of film or literature, for example, the game player is "embodied" in the game character, albeit in a way that only credits the game character with the status of a vehicle or equipment. In doing this, Newman highlights the complexity of the relationship between the game player and the game character, whom the player might interchangeably refer to as "him" and "me."³ This in turn draws attention to the degree to which the idea that playing a video game allows the player to be the protagonist of the adventure or "enter into" the fabricated world of the game—so

central to the discourse of escapism and immersion that underlies both the utopian or marketing-oriented accounts of video games and those that see them as threatening to cause addiction or modification of real-world behavior—is based on a naive simplification of the complex and dynamic regimes of vision and identification created by computer representations. At the same time, while Newman’s critique of simplistic accounts of representation and identification in video games is well founded, he perhaps overplays his hand when he argues that this critique compromises “the notion of identification and empathy in the primary-player–character relationship and, consequently, the privileging of the visual and of representation-oriented approaches” (see Klevjer, 2006, pp. 61–65).

Newman’s claim that the video game character is just a kind of equipment utilized by the player to act on the game world through a relationship of “vehicular embodiment” redeploys the logic of Cartesian dualism, with the player taking the role of disembodied *cogito* using the game character to act upon the digital *res extensa* of the game world. However, as with Cartesian dualism, the greatest weakness of such an approach is its inability to provide a satisfactory account of how these two components articulate with one another. The video game player cannot simply have an experience of “being” the game character that obliterates her experience of her own physical body, but if the game character were experienced as nothing more than a tool used to solve problems in the game world, gaming would surely not be as engaging for players as it is.

Furthermore, Newman’s argument against a focus on representation depends upon an implicit contrast between the aesthetics of video games and those of older visual media—paradigmatically film—that exaggerates their difference. The argument that an approach centered on representation is not useful to analyses of video games is supported by the claim that, rather than a stable and consistent regime of representation and identification, games present “multiple and apparently contradictory presentations of the self” (Newman, 2002, ¶ 15). However, in reality, the inverse is true: The regime of spatial representation employed by video games is generally much more unified, rigid, and consistent than that of film. Take Newman’s example of the arcade car-racing game:

So, in a CoinOp driving game, for example, it is possible to be, at once, seated in a mock-up car chassis, grasping a steering wheel with pedals beneath one’s feet, staring at a screen (presumably the windscreen), through which we view a remote, clearly mediated vision of ourselves as relayed by a camera in a trailing helicopter. (Newman, 2002, ¶ 15)

However, if we remove the mock-up car chassis—which would be an impractical addition not only to movie-going, but most video gaming as well—there is nothing about such a representational regime that conflicts with that of film. The contradictory logic of driving the car while watching the car from outside also does not seem so different from the logic of representation and identification we experience while

watching a car chase in a Hollywood film. While we do not control the car in the Hollywood film, we identify with the driver, and perhaps flinch at a near collision as if we were physically located inside the car, even as we watch the chase largely from a viewpoint outside the car. We are likely to also see shots depicting the interior of the car, and perhaps even the driver's viewpoint, but this only illustrates the degree to which representation and identification is more "multiple and apparently contradictory" in film than in video games. The film camera jumps from one location to another with no regard for spatial limitations: We are watching the driver of the car as if a passenger in the car, then we *are* the driver, looking out through the windshield as we drive, then we are suspended in the air above the chase looking down—a video game that defied common-sense notions of spatial relationship and viewpoint in such a way and at such speed would be so disorientating as to be impossible to play. In a typical shot–reverse shot sequence, we are placed in the position of the film's protagonist, only to be looking at the protagonist from the viewpoint of another character an instant later and then switching back again; the conventions of representation in video games are—and must be—far more constrained than this, for the simple reason that the player must be able to understand and navigate the virtual space of the game in a way unnecessary for the film viewer.

Positioning the Player

A problem with many existing accounts of the relationship between game player and video game is that they seek to resurrect pathologizing psychoanalytic descriptions of film viewers from the 1970s, and it is perhaps this tired and problematic approach that Newman is specifically seeking to reject. Games are no more able to subordinate or replace the embodied experience of their audience than any other media form; also, like other media forms, they could not function as media if they did. At the same time, however, the player seemingly has the experience of acting through a virtual body upon a virtual space—some kind of direct, preconscious engagement with virtual body and space is necessary in order to play the game, and without it games would presumably not be as viscerally enjoyable.

The lack of clarity and consistency in how the player's body engages with a game body does not mean that such engagement is of little value or relevance to the video game experience; it is difficult to clearly express the nature of my relationship with my own body, but I remain confident that this relationship is of crucial importance to an understanding of who I am and how I come to feel part of my environment. Rather than being thrown up by video games, questions about how we sense our own and other bodies, and how features of our environment can come to feel like a part of us, have already been raised by the nature of physical embodiment more generally, and the difficulty of answering them with regard to video games should come as no surprise given that no universally accepted answer has yet been formulated with regard to the rest of human experience. The addition of a game body to the equation, however, can perhaps make some aspects of this problem a little clearer, as player

body and game body can be isolated from one another more easily than *cogito* and *machina carnis*.

When I am controlling a vehicle or wielding a tool in the real world, these artifacts can produce new kinds of sensory stimulation for me and can even perhaps alter the boundaries of my experiential self (see Black, 2014), but 3-D virtual environments seek to create an experience of a fabricated, simulated space, and to do this, they must create the illusion that the boundaries of our experiential selves extend into that fabricated space (see Calleja, 2007, pp. 254–255). To do this, they most importantly seek to create a viewpoint situated inside the simulated space, and this viewpoint is in most cases attached to the virtual body of a game character. As will be discussed further subsequently, this viewpoint never is that of the game character in some straightforward way and is not dependent upon our crediting the game character with any particular psychological reality; but the fact remains that our sense of involvement with 3-D games, our sense of an ability to act on and in the virtual environment, is dependent upon the successful creation of this viewpoint. As a result, and *contra* Newman, we can say that in 3-D games, at least, a concern with the visual and representational, and the relationship between player and game character, is of crucial importance to understanding the experience of video game play.

When the unseen body of the protagonist Chell falls a huge distance in the game *Portal 2* (2011; as it regularly does), I, seeing the vertiginous drop seemingly through Chell's eyes, experience the kind of pleasant/unpleasant spasm of the stomach and tingling in hands and feet generated by a roller coaster, even though my body of course has not moved at all. When an FPS player ducks her head as a monster swoops down from a virtual sky, her relationship with the game character is clearly not one of user and equipment. The player is experiencing some kind of visceral, affective relationship with the game character, despite the fact that the game character cannot even be seen by her in many cases. There is a kind of empathy and identification going on here, and the inconsistency of perspective and representational regime does not nullify this relationship; it only demonstrates that the mechanisms and nature of this empathy and identification are more complex than naive accounts of "immersion" would suggest.

Perspective, Viewpoint, and Point of View; Ideal, Actual, and Avatar Viewpoints

It may seem self-evident that the first-person point of view is the most engaging by virtue of being the most realistic, and the most realistic by virtue of being the closest to real-world sensory experience. After all, in our everyday, embodied activity, we see the world from a viewpoint originating at our eyes, and our own bodies are largely invisible to us by virtue of being the point at which our sensory experience originates rather than being the object of that experience, and it is precisely this position that the first-person game seeks to reproduce. The development of "virtual reality" headsets such as the Oculus Rift is motivated by a belief that a virtual

environment will seem most realistic if we view it through a convincing first-person perspective, and the very terms “first-person” and “third-person,” with their implied relationships with “me” as opposed to “him” or “her” reflects this common-sense belief. Gamers interviewed by Kristine Jørgensen described themselves as “merging” with the protagonists of FPSs (2009, pp. 2–3), and Chris Chesher claims that “[f]irst person games generate a vertiginous sense of movement, and use sound and rapid movement within an enhanced depth-perspective space to give players a sense of visceral immersion” (2004, ¶ 18).

However, there are obvious problems with the first-person game’s efforts to simulate direct experience. At least when delivered through media such as screens and game controllers, mice, keyboards, and so on, the result is still a doubling, rather than convergence, of positions. For a player focused on a game, her field of view beyond the screen, sounds originating outside the gameworld, or the manipulations of a game controller may fade from consciousness, but they are still a part of her experience at some level. The FPS player’s body might seem to occupy the same spatial location as the game’s virtual protagonist, seeing and hearing the world from the position of a largely unseen virtual body, but the fit is only approximate. The player’s body remains outside the virtual environment, where it can see, hear, and feel things going on outside the game world; furthermore, while the player might produce actions that originate from the location of the virtual protagonist, she does so by means of a kinesthetic performance enacted by her physical body outside the game world, a performance distinctly different from the virtual actions it generates. The little that can be seen of the virtual protagonist—for example, the hand seating a magazine in a gun—only highlights the nonidentity of the location and actions of the player’s and game character’s bodies.

As noted earlier, fluidity or inconsistency of viewpoint is hardly peculiar to video games. Point of view can shift in novels, and in film it is routine for viewpoint to shift with no regard for the physical properties of the filmic environment (looking down from the sky, through windows, and even taking on the first-person viewpoint of people or even animals or inanimate objects), yet the ability of these media to produce identification and empathy is taken for granted. A key difference in the FPS is that it presents a simulated perspective that is rendered in real time to reproduce the laws of Euclidean space, and this is key to its sense of realism: The viewpoint seems to be fixed inside the representational world and obey its physical laws as if it were physically manifested there (e.g., the viewpoint does not constantly jump from shot to reverse shot or cut away to a different location as it might in a film).⁴ However, even in an unadulterated first-person viewpoint, the reality of how this works is complex and multiple. This is immediately obvious if we think about the relationship between the player’s viewpoint and the viewpoint of the game itself; the representational regime works by creating a sense that there is no difference between the two, but of course there is. If we try to evaluate the game’s realism using straightforward spatial relationships, it immediately breaks down; for example, Newman describes the relationship between player and game character as one of “vehicular

embodiment,” but what can this mean in this context? How can the player be “inside” the game character, as suggested by this term? How can player and character have any spatial relationship to one another at all, given that they exist in two fundamentally incompatible sets of spatial coordinates: the real-world space of the player’s body and the simulated space of the game?

Any attempt to create a realistic visual simulation of space is complex and contradictory, whether it be in a Renaissance painting, a Victorian optical toy, or a virtual environment. In order to be a success, however, such a simulation must seem, on the surface, to simply recreate our everyday visual experience, and this can obscure the complexity and inconsistency of its operation. As a result, we need to spend some time considering the strategies used by video games to simulate space visually, and the process by which they have appropriated and adapted these strategies from older representational technologies.

In order to do this, it is important to initially differentiate my use of the terms *perspective*, *viewpoint*, and *point of view*. The general usage of these terms often overlaps or doubles up, but for the purposes of the following discussion it is necessary to draw hard distinctions between them in order to avoid confusion.

I will use the terms *perspective* and *linear perspective* interchangeably to refer to the arrangement of features within a 2-D image or images in order to simulate depth. In other words, whether it be in a Renaissance painting or a video game, it is the creation of an illusion that the viewer can see “into” a simulated space beyond the flat plane of the image (Kemp, 1990, p. 342).

The term *viewpoint* will refer to the camera’s position in space, from which perspective originates, while *point of view* will be used to refer to the coincidence of camera viewpoint and a game character’s (simulated) viewing position. So, in an FPS, the player is presented with the game protagonist’s point of view, an effect created by situating the viewpoint in the protagonist’s head.

This usage has been adapted from cinematic conventions, and when they refer to camera position it may be objected that video games, unlike films, do not involve the use of a camera. However, camera position is fundamentally important to the depiction of any computer-simulated 3-D space such that, while no actual camera is used to produce such representations, they all feature a “virtual camera.” With no camera (actual or virtual), there would be no viewpoint, and without a viewpoint there would be no perspective. Without perspective, there can be no visual simulation of 3-D space. While point of view is far more common in games than film, it remains an optional component of the representation; viewpoint, on the other hand, is necessary for any perspectival representation to be intelligible to a viewer.

Perspectival representations work by confusing the distinction between these different components in order to produce the illusion of sharing the location of the camera or character within the represented space and thus of being within the representational space ourselves. By pulling apart these three different components and looking at the functions they serve in creating the game player’s sense of moving around “inside” a simulated space, we can develop an account of games’

representational regime sophisticated enough to allow an investigation of how, and to what extent, the player has a genuine experience of being embodied in the game world. I will begin with a discussion of how perspective operates within the 3-D video game.

Since Brunelleschi “invented” linear perspective in the 15th century (Kemp, 2006, p. 15), realistic visual representations of space have been understood to be most importantly produced through the mechanism of a fixed, ideal viewing position outside the image. Linear perspective creates the illusion of space in realist painting, film, and the 3-D computer graphics of video games (see Manovich, 2001, p. 184).

However, while the development of rendered 3-D graphics is clearly part of a long history of technologically simulating perspective, this does not mean that they represent nothing more than a continuation or refining of an existing mode of representing and looking. The relationship between user and computer-generated perspectival image is different from its pre-digital progenitors in important ways (cf. Taylor, 2003).

Linear perspective depends, of course, on the creation of a vanishing point inside the image, which produces the effect of a monocular viewpoint situated in a singular spatial position relative to the simulated space of the image. It therefore creates twinned points at either end of a line of sight, both of which seem to float outside the 2-D plane of the image: the vanishing point that is *through* the image, terminating the line of perspective inside its illusionary space—the point of invisibility on which the representational landscape converges—and the viewpoint that is *outside* the image, the point at which the field of vision of the viewer should originate. This second point has only an abstract, mathematical reality until photography, when it comes to originate with the actual physical presence of the camera (although of course it remains attached to the image after the camera has gone). When these two invisible points are lined up—in other words, when the viewer places herself at the viewpoint outside the image so that the lines of the painting fall away toward the vanishing point inside it in a way that conforms to the effect generated by 3-D space—the viewer experiences an illusion of depth. When a painting is created, an invisible, phantasmal viewer is created with it, hovering directly in front of the painting a certain distance from its surface, and it is only by “becoming” that viewer, positioning ourselves so that we are inhabiting the same point in space as the phantasmal viewer, that we experience the effect of perspectival realism. We therefore need to differentiate between two kinds of viewpoint: the *actual viewpoint*, that is, that of the living body engaged in the act of looking from a particular physical location, and the *ideal viewpoint*, a purely theoretical viewing position created as a by-product of linear perspective, with which the actual viewpoint must coincide⁵ in order for the simulation of space to succeed.

This technique is initially developed in painting and drawing, but subsequent technologies that mechanically reproduce perspective work in the same way. The key development with film is that this ideal viewpoint becomes mobile. The camera is able to record shifts in its viewpoint over time, creating a sense in the viewer that

she is moving around inside the representational space of the film. The realism of 3-D video game images clearly functions in the same fashion, which is not surprising given video games' aesthetic debt to cinema. The rendering of "3-D" environments in games is effected by simulating the trajectories of the rays of light that are captured by a film camera, mathematically plotting the lines of simulated light that will produce the effect of perspective. Of course, such 3-D graphics are no more 3-D than movies: They are still produced as flat, 2-D renderings of space. What makes them seem "more" 3-D than cinema is that video game technology brings the capacity to manipulate the vanishing point of the image, whose movement causes an equivalent change in the ideal viewpoint attached to it. In effect, this means that the player can swivel the ideal viewpoint from side to side by manipulating a controller. Because the player cohabits the space of the ideal viewpoint, this in turn creates a sense that the player is moving around with it. This is entirely illusory, of course: Not only does the player's body not move, it also doesn't change its actual viewpoint relative to the image—it is only the relationship between the image and the ideal viewpoint that changes as the image projects differing points of perspectival convergence outside itself in real time. If a player has a sense of traveling forward into the 3-D space of the game world, it is not because the player's actual viewpoint has changed—the player's body is in the same position it was in before, and her viewpoint relative to the screen is no different. Rather, the perspective constructed around the ideal viewpoint has shifted forward, and as long as the player identifies with that ideal viewpoint, and feels that her perspective is meshed with that of the ideal viewpoint, then when the ideal viewpoint shifts perspective, the player will feel that her body is shifting along with it.

What the video game adds to film's introduction of a mobile ideal viewpoint is an experience of, not simply moving through the representational space, but of being able to physically act upon that physical space (see Calleja, 2007, pp. 254–255; Klevjer, 2013, p. 7). It does this by creating a sense that the ideal viewpoint is tied to a physical agent within the simulated space, usually a simulated body. The game's creation of perspective therefore adds a third viewpoint to the two already discussed in relation to painting and film, one that is *inside* the representational landscape. With the perspectival painting, the ideal viewpoint is constructed at a fixed point, and the actual viewpoint can then occupy the same space as the ideal viewpoint. With film, the ideal viewpoint position is mobile, engendering in the actual viewpoint a sense of movement, but that ideal viewpoint is still physically disconnected from the space of the film itself: It is not tethered to anything inside the image, being able to switch between an infinite variety of positions relative to the features of the represented environment. In video games, this ideal viewpoint becomes physically manifested in the represented environment in some way. And in games focused on a single protagonist, this ideal viewpoint enters into a relationship with the posited viewpoint of a game character in order to create what I will refer to as the *avatar viewpoint*. Furthermore, in the FPS, the ideal and avatar viewpoints occupy the same

spatial coordinates in order to create the sense of a simultaneity of, not only ideal and actual viewpoints but the viewpoints of player and game character.

The sense of agency created by this effect both heightens and obscures the relationship between the living body of the player and the phantasmal body of the ideal viewpoint that shares the same point in space. The ability to move the viewpoint through the representational space increases the illusion that there is only one viewing position—that of the player—that generates perspective, but at the same time other factors—for example, the fact that the player cannot see her body in the space of the game—highlights the differences between them.

Such complications notwithstanding, the creation of a convincing illusion that the perspectives of the ideal and actual viewpoints are perfectly meshed is a key tool used in games to create a sense of visual realism. In the FPS, the game creates a triple convergence between one real and two fictitious viewers: The actual viewpoint of the player, the ideal viewpoint outside the image, *and* the avatar viewpoint inside the game are all aligned in such a way as to create the illusion that they inhabit a single point in space, producing a sense that, not only is the player's viewpoint moving around inside an illusory 3-D space, but also that this viewpoint is being shared with that of an illusory body inside that space.

And yet, despite this fact, there are many examples of games that intentionally break up or complicate this stacking effect by situating the ideal viewpoint a short distance away from the avatar viewpoint. In third-person games, the ideal and actual viewpoints remain anchored to a feature of the representational landscape, but the avatar viewpoint is not directly superimposed over the ideal and actual viewpoints; rather, it is situated a short (but usually fixed and stable) distance away from the ideal and actual viewpoints, leaving them trailing around behind the game character's shoulder like a balloon on a string.

The third-person game should immediately raise a question about the role of viewpoint in games more broadly. If a video game is seeking to create a sense of being a part of the virtual environment and acting upon it, isn't a third-person representation innately inferior to a first-person one? In a first-person game, the stacking of ideal, actual, and avatar viewpoints creates the illusion that I am the hero moving through and acting on the game world, but wouldn't seeing the hero as a separate entity situated a short distance away from my simulated perspective break this illusion? It seems common sense that the third-person viewpoint is "more detached than the first person view" (Chesher, 2004, ¶ 15); if this is true, when given a choice between first-person and third-person perspectives, why would game designers opt for the latter, and why wouldn't players experience the latter as fundamentally less immersive and involving than the former?⁶

The common and successful use of third-person representations should, in itself, establish that the player's relationship with the game character is not simply one of direct identification. If it was, third-person games would always be less involving than first-person games.⁷ Also, by contrast, in other audiovisual media such as film and television, it is standard practice for the ideal viewpoint to be situated outside

any particular body, and first-person point of view is rare, and fleeting when it does occur. Is their greater use of point of view the reason why video games supposedly produce a greater level of “identification” than older audiovisual media (Shaw, 2010, pp. 147–148)? If so, why haven’t these older audiovisual media simply adopted a more extensive use of point of view themselves? The technologies used to create and display these media forms are just as capable of doing so as video games.

Film Bodies and Game Bodies

When games are heroized for putting the player inside a simulated world or decried for allowing players to act out violent fantasies or when the nature of the player’s identification with the game character (or lack thereof) is being debated, the implied nature of the identification is quite straightforward: In effect, there is the player, and there is the game character, whom the player effectively becomes while playing the game. But, in reality, how the player is positioned relative to the game world, and which vantage point on it the player is invited to identify with, is much more complex. The actual, ideal, and avatar viewpoints are separate and distinct in their attributes and, while the representational regime of the game seeks to create a sense that they have all been blended into a single coherent unity, they remain separate from one another, superimposed one upon another but quite unstable and vulnerable to dislocation.

This should come as no surprise given the existing investigations of identification as it occurs in film, a medium whose greater age has allowed for a more thorough theorization than video games. As I have noted previously, the fundamentals of this representational regime are more apparent in film; rather than video games being more inconsistent or unstable in their regimes of representation and identification, they are actually differentiated from film by their more rigid and limited organization. At the same time, the ability to manipulate viewpoint in games introduces an additional, different kind of complexity, and it is the greater potential for instability introduced by this that necessitates a more stable and limited representational regime in order for the video game’s simulation of space to be intelligible to the player.

Vivian Sobchack’s (1992) phenomenology of film spectatorship contains a refutation of the idea that the experience of film viewing is derived from a single, unified experience of vision. Where psychoanalytic accounts of spectatorship have cast the viewer as subordinated to the viewpoint of the camera—the viewer eagerly swapping her own imperfect subjectivity for “the fantasy . . . of a ‘transcendental subject’” (Iversen, 2005, pp. 194–195)—Sobchack highlights the impossibility of the viewing body ever being suppressed or obliterated by the machinery of film. Technologies of vision have more widely and repeatedly been cast as replacing, or even transparently extending or transforming, human vision, but such descriptions ignore the fact that the images produced by such technologies are always, themselves, objects of human vision, rather than replacements for it. As Sobchack notes, when

watching a film, “we can see the seeing as well as the seen, hear the hearing as well as the heard, and feel the movement as well as see the moved” (1992, pp. 10–11). The camera/screen assemblage does not replace the eye and mind of the viewer; it presents another experience of viewing, which then itself becomes the object of the viewer’s vision (Sobchack, 1992, p. 141).

Sobchack describes the viewer’s relationship with the film as related to the subject’s visual engagement with other viewing bodies; however, where this latter engagement is one in which we are presented with the physical presence of another body whose viewing subjectivity is hidden from us, film presents the viewing subjectivity of another whose body is hidden from us: “Unlike other viewing persons I encounter, the film visibly duplicates the act of viewing from ‘within’—that is, the introceptive and intrasubjective side of vision” (Sobchack, 1992, pp. 137–138). This leads Sobchack to give an account of a “film body,” that is, the viewing subjectivity that the technology of film creates between the film maker and film viewer.

Sobchack’s account of the film body is particularly useful for the current discussion, given that games that utilize an avatar viewpoint explicitly seek to produce an equivalence between the player’s point of view and that of the game character. This does not mean that the body of the game character is equivalent to Sobchack’s film body—it is not—but the additional components and complexities introduced to Sobchack’s account by the avatar viewpoint allow a greater understanding of how the game player relates visually to the game’s virtual environment.⁸

Following Sobchack, we can maintain that the three different viewpoints—actual, ideal, and avatar—while they can blend into one another in certain ways, are never a unified, homogenous whole. The player’s visual relationship with the game world is therefore fluid and multiple and is further complicated by the nature of the player’s capacity to act upon the game world. The player’s visual relationship with the world of the game is determined by the ideal viewpoint of the virtual camera, but the player’s capacity to act upon the virtual environment is tied directly to the game body—regardless of whether the ideal viewpoint and game body occupy the same point in the representational space. As a result and as illustrated by the third-person game, the player can be looking at the game body as a separate, externalized entity while still feeling that her capacity to act on and in the game is expressed by that separate, externalized entity.

Furthermore, a more explicit contrast between film and video game can be found elsewhere in Sobchack’s characterization of the film body. In seeking to establish the particularity of the film body and the visual experience that it makes available to the viewer and its separateness from the points of view of the characters within a given film, she cites as evidence the fact that not only does the viewpoint of the camera rarely take on the point of view of any of the characters in a film but doing this seemingly renders the film’s representational regime less viable. Sobchack’s point, therefore, is that the film body has its own attributes that are quite different from living bodies, and which therefore seem jarring and strange if they are attributed to a living body; but it also raises the question of why, if Sobchack’s claim

about the irreconcilability of film body and first-person point of view is justified, the use of first-person point of view—in other words, the conflation of ideal viewpoint and avatar viewpoint—is a commonplace in video games.

In support of her claim for the awkwardness of first-person point of view in film, Sobchack cites what is considered the classic illustration of the nonviability of sustained first-person film narrative: the 1947 film noir *Lady in the Lake*, in which for almost the entire film the character of Philip Marlowe is positioned in the representational space in a way now familiar to players of first-person video games (see Rehak, 2003, pp. 119–121).

The protagonist and perceptual autobiographer, detective Philip Marlowe, is predominantly visible only *in* the perceptual correlations of “his” (the film’s) vision in the way that we appear materially visible to ourselves in our visual perception. That is, Marlowe sees himself *as* he sees only through his *visible reflection* in mirrors or other reflective surfaces. As well, he sees himself as directly and materially visible only in those *parts* of his body that are brought before his eyes—when, for example, the perceptual correlation makes visible a hand brought up to light a cigarette that hangs suspended from unseen lips. Otherwise, Marlowe is invisible to himself and to us but nonetheless constantly implicated as a *physically material* and *human presence* enabling the visible perception. (Sobchack, 1992, p. 230, emphasis in original)

The “strange discomfort, alienation, and disbelief experienced by the film’s spectator” (Sobchack, 1992, p. 231) are taken as proof of the fundamental nonviability of this representational scheme, and the causes of this discomfort have since been discussed at great length by numerous authors. But if watching a film with an almost uninterrupted first-person point of view is so off-putting, why is it that many of the most successful video games of all time have utilized the same representational scheme without issue?

Alexander R. Galloway has sought to provide an explanation for the different status of point of view in film and video games and furthermore has resisted the temptation to psychoanalyze viewers and players in order to do so (2006, Chapter 2). Galloway suggests that the subjective point-of-view shot, rather than working to make the viewer feel that her subjectivity and that of a character are the same, has been most successful in film when alienating us from the character whose subjectivity we are being presented with, and, furthermore, these alien subjectivities are often computerized in some way, suggesting an affinity with technologized vision that makes it more appropriate for a video game than a film:

[T]he merging of camera and character in the subjective shot is more successful if the character in question is marked as computerized in some way. The first-person subjective perspective must be instigated by a character who is already mediated through some type of informatic artifice. Necessary for this effect are all the traces of computer image processing: scan lines, data printouts, target crosshairs, the low resolution of

video, feedback, and so on. In other words, a deviation from the classical model of representation is necessary via the use of technological manipulation of the image— a technological patina. (Galloway, 2006, p. 56)

This idea of an evolution from the cyborg vision of the Terminator or Robocop to that of the FPS is an evocative one, but it fails to account for several aspects of the film–video game contrast in point-of-view representations. While Galloway gives an extensive listing of the successful use of point-of-view shots in films, not all of them represent computerized vision and none of them succeed where *Lady in the Lake* failed. That is, computerized vision is only a subset of a larger category of “alienated, disoriented, or predatory vision” (Galloway, 2006, pp. 68–69) that also includes monsters and murderers, the drugged and insane, and their visual points of view are only presented fleetingly, rather than being the text’s primary mode of representation as is the case with both *Lady in the Lake* and the FPS (see Brooker, 2009, pp. 127–128).

Galloway does note the effectiveness of point of view in creating a sense of movement through space, however, and I think this suggests a slightly different explanation. Among other perspectives, Sobchack refers with (qualified) approval to an explanation for *Lady in the Lake*’s failure presented in the 1960s by the French film theorist Jean Mitry:

He sees the failure of *Lady in the Lake* to convince the spectator that s/he is Philip Marlowe as a failure based not so much on the invisibility of the character’s body as on real *bodily difference*. Mitry emphasizes the difference between the spectator’s body sitting relatively quiescent in a theater seat and the film’s body invisibly living out, through the activity of the camera, a kinetic life and activity clearly not shared bodily by the spectator . . . [A]lthough we, as spectators, may be sympathetic to cinematic perception and, indeed, may intentionally parallel the film’s and/or character’s bodily position and perceptual bias as it intends toward and inhabits a world, we physically and materially occupy our *own* bodies and space. The perception whose intentional interest we share belongs always to *another* perceiving and embodied subject, no matter how introceptively it is visibly presented as visual for us. (Sobchack, 1992, pp. 233–234, emphasis in original)

Of course, the same could be said for the video game player, who is not walking, jumping, or shooting as the game character does these things. At the same time, however, I think it draws attention to a more general contrast between the levels of activity and agency in film and video game. The video game produces an experience of being in and acting upon a simulated space distinctively different from film.

Agency and Movement

While the video game player to some degree and in some fashion shares the experience of action in the video game world, the fact remains that the video game player

can never be claimed to have an experience of simply being the video game body acting upon the virtual environment in which that body is located. The player always is—and always must be—an embodied physical presence who does not look out onto the virtual environment through the game character's eyes but looks at a screen with her own eyes and does not engage in embodied activity in the virtual environment but rather manipulates one or more control devices such as a game controller, keyboard, or mouse. There is a degree of transparency to the manipulation of control devices that comes with skilled use, but this manipulation remains a part of the player's experience, and constitutes an embodied experience of movement quite different from the bodily actions being represented in the game.

However, while the physical actions of the player and the represented actions of the game body might be quite different, the fact remains that the player is engaged in physical activity in a way that the film viewer, for example, is not, and furthermore—and crucially—that action is synchronized with the actions of the game body. That is, not only is the game player manipulating a control device while looking at the imagery of the game as opposed to the film viewer who is “sitting relatively quiescent in a theater seat,” but the activity of the game body is instantly responding to those manipulations. Human bodies are never completely passive or inactive, but if a film viewer shifts in her seat or stuffs a handful of popcorn into her mouth, this produces no effect in the images she sees.

I would argue that this disconnection between the embodied actions of the film viewer and the film's images facilitates the relationship between the film viewer and Sobchack's film body; there is no sense that the viewer's body and the film body are directly equivalent, which would create a sense of dislocation or awkwardness where the visual experience of the film body follows a logic alien to embodied human perception. For example, if the viewer felt that she was directing or controlling the vision of the film body, common features of film such as the shot–reverse shot would presumably be jarring and disorientating, given that they are utterly alien to the attributes of embodied human vision. The other side of this, however, is that the use of a first-person point of view in film feels unpleasant and constricting; when the film body attempts to impersonate a living human body, its obvious lack of attributes fundamental to the viewer's embodied perception—such as a sense of agency and an intentional relationship with the environment—creates a sense of discomfort. Sobchack describes the experience of watching *Lady in the Lake* as one of claustrophobic constriction, and presumably this comes from the sense of being trapped inside a body over which one has no control; the viewer can only peer helplessly out of its eyes as it is propelled through the world by an alien subjectivity whose workings are hidden from its passenger. This explains the fact that point-of-view shots in film are largely restricted to depictions of “alienated, disoriented, or predatory vision”; these are subjectivities that, rather than inviting an expectation of familiarity and control in the viewer, have either had their capacity to see or act deranged or are driven forward by a kind of monomaniacal automatism—these gun-toting cyborgs or psycho killers are driven to disregard all other possible actions or objects of attention as they

single-mindedly hunt down their prey. Unlike Philip Marlowe in *Lady in the Lake*, these figures do not seem capable of distraction or shifts of interest; the viewer is not frustrated by her inability to control these bodies because they are bodies that are understood to be unable even to consciously control themselves.

In the video game, on the other hand, the player has a sense of agency in the simulated space. In a first-person game, if an object in the virtual environment catches the player's eye, she can center the ideal viewpoint on it, can move that viewpoint through the virtual space toward it, and can possibly interact with it. The fact that the viewpoint being centered on the object is not the embodied viewpoint of the player but a viewpoint generated by the game, or that the player is not physically walking, shooting, or picking up the object, is not crucial here; what is crucial is the experience of agency and intentionality. The movements of the game body are not the same as those of the player body just as the viewpoint on the game world is not the same as the viewpoint of the player, but the movements of the player and game character are synchronous and seemingly generated by the player's intentionality. This is presumably why even games with quite rudimentary, low resolution images can create a sense of involvement: The key factor is not a simultaneity of character point of view and player viewpoint—and in fact such a simultaneity can only ever be imperfect and unstable—but rather a simultaneity of the activity represented in the game and the actions of the player's body, which are generated in response to that activity in a circular fashion. This provides an answer to the question raised about third-person games: Why does the separation of player and character viewpoints not lessen the player's sense of involvement? First, even in a first-person game, the viewpoints of player and game character are never truly unified, meaning that the difference between first-person and third-person representation is only one of degree, and, second, a more important kind of simultaneity between player and game character—a simultaneity of action rather than viewpoint—is just as present in the third-person game as the first-person. In fact, given that the player has a greater awareness of the game body and its actions when the ideal viewpoint is situated outside the game body, it is possible that—at least in some instances—the third-person game creates a greater sense of simultaneity of action than the first-person.

Shared Embodiment

There is certainly reason to believe that a simultaneity of action and perception can change our experience of where our bodies are and what they are doing. Just as the film viewer's visual experience becomes a synthesis of viewing body and film body, rather than either a replacement of viewing body by film body or a clinical, disengaged inspection of the screen, so the game player experiences a coming together of playing body and game body. As noted earlier, if this did not happen, video games would presumably not be as successful as they are. Playing a video game does not feel like typing a letter on a word processor or dragging a file from a virtual folder to a virtual trashcan. Just because the game player does not wink out of existence as an

identifiable subjectivity or locus of sensory experience while playing a game does not mean that her embodied experience does not on some level incorporate the game body.

Research into how we experience our own living bodies and understand those of others demonstrates how natural and basic such a relationship can be. Neuroscientist H. Henrik Ehrsson has become well known for experiments that cause subjects to believe that part or all of another body—living or artificial—belongs to them. In the most famous of these experiments, a team lead by Ehrsson fitted each subject with a “virtual reality” headset that fed her images from cameras attached to the head of another person standing behind her. In other words, the experiment set up a visual relationship much the same as that between FPS player and game body but with a real living body substituting for the virtual body of the game character.⁹ When this was combined with tactile stimulation widely used in body-ownership experiments, subjects could be made to feel that they were having an “out-of-body experience,” looking at their own bodies from outside themselves or, in a variation on the format, even feel that their bodies had been replaced by a mannequin or doll (Guterstam & Ehrsson, 2012; Petkova & Ehrsson, 2008; van der Hoort, Guterstam, & Ehrsson, 2011).

Because the subjects’ movements were manifested by the body being looked at, the overall effect was not like that of playing a first-person game, but rather of playing a third-person game. The loci of viewpoint and action were disarticulated, seemingly occupying different points in space so that the subject’s capacity for action was exercised by the body being seen in front of her rather than by the one she seemed to be watching it with. Our capacity to have such an experience makes a close identification between game player and game body in third-person games seem more natural and attainable than it might otherwise appear.

Significant for the earlier discussion of simultaneity of action and perception when successfully employing third-person viewpoint in video games, a key component of all such body-ownership illusions, from the originary “rubber hand illusion” onward (see de Vignemont, 2011), has been the subject’s observation of tactile stimulation of the surrogate body that is synchronous with the feeling of tactile stimulation of her real body. Again, it is the perceived simultaneity of what the subject sees happening to the simulated body and what the subject does or feels with her own body that creates a sense of commonality.

Of course, it would be naive and absurd to suggest that, when playing a video game, the player is reproducing the experiences documented in such experiments. In fact, it would be antithetical to my argument, as these experiments claim to be doing what simplistic accounts suggest video games do: merging player and character into a subjective unity. The subjects of these experiments felt that their subjectivities had been transferred into another body, and the elaborate procedures required to produce such an effect only highlight the absurdity of suggesting that playing a video game could produce something equivalent. But such phenomena do draw attention to the fact that my sense of what is and isn’t my body, of where my body is, and how

far my body extends is always quite fluid. A blind person senses her environment through the end of a cane; someone else will flinch while watching a third person fall over as if afraid that any resulting injury will be manifested on her own body. We routinely shift the boundaries of our bodily experience in multiple ways and to multiple degrees, and the simulated spaces and bodies of video games generate sensory, kinesthetic, and affective engagement by inviting us to extend this capacity into their virtual worlds.

Conclusion

Video games have demonstrated an ability to create a sense of immediacy and involvement in players, which suggests that a sense of involvement in their represented events and agency in their simulated environments is a key part of their appeal. At the same time, however, to suggest that players experience games in the same way that they experience embodied everyday activity or that the simulated characters and events of games can swamp their existing subjectivity or embodied experiences is implausible. Like any media form, the video game can create new experiences by producing another layer of embodied experience that is able to articulate with the foundation of embodied experience that is with us all the time, creating novel combinations. Rather than confusing or replacing our everyday sense of where our bodies are or how they can sense or act upon our world, video games provide an experience of vision and action that is multiple and distributed across physical and simulated space. It is not unique in doing so, but, while our bodies are capable of having such experiences without them, interactive computer-generated simulations are well suited to both producing and investigating experiences of this kind.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

Notes

1. See, for example, Behrenshausen (2007, p. 336).
2. Throughout this essay, I use the term “immersive” as representative of naive, simplistic understandings of the relationship between player and game that suggest simulated digital environments can somehow replace the player’s embodied, material reality (see Salen and Zimmerman on the “immersive fallacy” (2004, pp. 450–451; see also Klevjer, 2006, pp. 65–68)).
3. For more on this phenomenon, see Hitchens, Drachen, and Richards (2012).

4. Most players of third-person games are likely to have had the experience of a virtual camera becoming wedged in some awkward or unhelpful location by an inconveniently located wall, for example, while such problems do not afflict films.
5. Coincide in location but not necessarily in subjectivity—see below.
6. There are clear reasons why a game designer might *want* to employ a third-person viewpoint; for example, it makes possible various kinds of spectacularization and commodification of game bodies that are useful particularly in role-playing and some kinds of action games. At the same time, however, it seems unlikely that these considerations would be enough to outweigh the disadvantage of a third-person position making the game less engaging for the player.
7. In fact, the research of Hitchens et al suggests that variation in levels of identification in video games is caused more by differences in players than differences in regimes of representation (Hitchens, Drachen, & Richards, 2012).
8. For another application of Sobchack's "film body" to video games, see Crick (2011).
9. See Cleland (2010, pp. 87–88) for another discussion of this experiment in the context of video games.

References

- Assassin's Creed*. (2007). Montreuil: Ubisoft.
- Behrenshausen, B. G. (2007). Toward a (Kin)aesthetic of video gaming. The case of dance dance revolution. *Games and Culture*, 2, 335–354.
- Black, D. (2014). Where bodies end and artefacts begin: Tools, machines and interfaces. *Body & Society*, 20, 31–60.
- Brooker, W. (2009). Camera-eye, CG-eye: Videogames and the "cinematic." *Cinema Journal*, 48, 122–128.
- Calleja, G. (2007). Digital game involvement a conceptual model. *Games and Culture*, 2, 236–260.
- Chesher, C. (2004). Neither gaze nor glance, but glaze: Relating to console game screens. *SCAN Journal of Media Arts and Culture*, 1, 98–117.
- Cleland, K. (2010). Prosthetic bodies and virtual cyborgs. *Second Nature: International Journal of Creative Media*, 2, 74–101.
- Crick, T. (2011). The game body: Toward a phenomenology of contemporary video gaming. *Games and Culture*, 6, 259–269.
- de Vignemont, F. (2011). Embodiment, ownership and disownership. *Consciousness and Cognition*, 20, 82–93.
- Galloway, A. R. (2006). *Gaming: Essays on algorithmic culture*. Minneapolis: University of Minnesota Press.
- Grand Theft Auto* (developed by Rockstar). (1997). New York, NY: Rockstar Games.
- Guterstam, A., & Ehrsson, H. H. (2012). Disowning one's seen real body during an out-of-body illusion. *Consciousness and Cognition*, 21, 1037–1042.
- Hitchens, M., Drachen, A., & Richards, D. (2012). An investigation of player to player character identification via personal pronouns. Presented at the 8th Australasian Conference on Interactive Entertainment: Playing the System, Auckland, NZ.

- Iversen, M. (2005). The discourse of perspective in the twentieth century: Panofsky, Damisch, Lacan. *Oxford Art Journal*, 28, 191–202.
- Jørgensen, K. (2009). “I’m overburdened!” An empirical study of the player, the avatar, and the gameworld. Proceedings of the 2009 DiGRA International Conference: Breaking New Ground: Innovation in Games, Play, Practice and Theory, Brunel University, London.
- Kemp, M. (1990). *The science of art*. New Haven, CT: Yale University Press.
- Kemp, M. (2006). *Seen/unseen: Art, science, and intuition from Leonardo to the Hubble telescope*. Oxford, England: Oxford University Press.
- Klevjer, R. (2006). *What is the avatar: Fiction and embodiment in avatar-based singleplayer computer games*. Unpublished PhD thesis, University of Bergen, Bergen.
- Klevjer, R. (2013). *Representation and virtuality in computer games*. Presented at the the Philosophy of Computer Games Conference, Bergen.
- Manovich, L. (2001). *The language of new media*. Cambridge, MA: The MIT Press.
- Newman, J. (2002). The myth of the ergodic videogame. *Game Studies*, 2.
- Petkova, V. I., & Ehrsson, H. H. (2008). If i were you: Perceptual illusion of body swapping. *PLoS ONE*, 3, e3832.
- Portal 2*. (2011). Bellevue: Valve corporation.
- Rehak, B. (2003). Playing at being: Psychoanalysis and the avatar. In M. J. P. Wolf & B. Perron (Eds.), *The video game theory reader* (pp. 103–127). London, England: Routledge.
- Salen, K., & Zimmerman, E. (2004). *Rules of play: Game design fundamentals*. Cambridge, MA: MIT Press.
- Shaw, A. (2010). *Identity, identification, and media representation in video game play: An audience reception study*. Unpublished PhD thesis, University of Pennsylvania, Philadelphia.
- Sobchack, V. (1992). *The address of the eye: A phenomenology of film experience*. Princeton, NJ: Princeton University Press.
- Taylor, L. (2003). When seams fall apart: Video game space and the player. *Game Studies*, 3. *Tomb Raider* (developed by Core Design). (1996). Eidos interactive.
- van der Hoort, B., Guterstam, A., & Ehrsson, H. H. (2011). Being barbie: The size of one’s own body determines the perceived size of the world. *Wimbledon: PLoS ONE*, 6, e20195.
- World of Warcraft*. (2004). Irvine: Blizzard entertainment.

Author Biography

Daniel Black lectures in the School of Media, Film and Journalism at Monash University, Australia. He is one of the editors of *Complicated Currents: Media Flows, Soft Power and East Asia*, the author of *Embodiment and Mechanisation: Reciprocal Understandings of Body and Machine from the Renaissance to the Present*, and has written a number of articles and book chapters on topics relating to our embodied relationship with technology.