Embodying the Avatar in Videogames

Mahdi Dibaiee
Student ID: 20200705

May 2022

1 Introduction

Videogames are a pervasive part of lives of children and adults alike, with 73% of Americans older than 2 years engaging with them (Group, 2019). Playing videogames can be seen as an activity that is done through our fingertips and with our visual apparatus focused on a screen, without involvement of the rest of our body, and it is usually considered as such from a cognitivist point of view (Campbell, 2012; Gee, 2003; Klimmt and Hartmann, 2006) however this raises the question of whether videogames can alternatively be thought of as an embodied experience, and if so, how can we formulate them as such, and what factors are at play?

Virtual reality videogames are more commonly studied from an embodied perspective, since they lend themselves to the framework more easily by being more engaging to the whole body and by the fact of their immersive experience, however the same question can be asked for non-virtual reality games, with keyboard and mouse or the controller, and the screen.

We will first talk about what do we mean by embodiment when we say playing videogames is an embodied experience, and this is a very important part of our discourse. We then continue to talk about what motivates us to think that videogames fit such notions of embodied experience, and from there we further ask questions about the factors at play, including, but not limited to, camera control and perspective and its relationship with peripersonal space and the social aspect of videogames.

2 Did Somebody Say Embodiment?

The question of whether we can think of playing videogames as an embodied experience is quite puzzling, and it requires unraveling questions that are unanswered about what embodiment means, how do we distinguish it from else, and how does something like playing videogames fit into this picture. There are different accounts of embodiment, and they stand in contrast to cognitive psychogolical accounts. Cognitive psychology accounts study mental processes, which are usually associated with the brain, and where the body is thought

of as an input and output interface with the world that is controlled by the brain. (Neisser, 2014; Anderson and Crawford, 1980). There are numerous accounts of embodied experience, we will review some of them and lay out our understanding of embodiment, one which allows us to discuss videogames in its light.

Thelen (2000) gives an account that focuses mostly on the fact that our experiences arise because we have a particular kind of body with particular capacities and apparatus that lead to us experiencing the world as we do. This might be one of the most high-level accounts that shares a considerable amount with most other embodiment accounts:

> "[T]o say that cognition is embodied means that it arises from bodily interactions with the world, from this point of view, cognition depends on the kinds of experiences that come from having a body with particular perceptual and motor capacities that are inseparably linked and that together form the matrix within which memory, emotion, language, and all other aspects of live are meshed."

With this account, it is necessary to consider the body as a constitutive part of cognition, not merely an input/output system controlled by the brain. Questions about cognition only make sense with consideration of the way we interact with the world with our bodies.

Merleau-Ponty's phenomenological account unifies the body and the mind and instead of talking about them separately, he proposes talking about an intentional, lived body, that is continuously adapting to the world through formation of habits:

> The body's orientation toward the world is essentially temporal, involving a dialectic between the present body (characterized, after Husserl, as an "I can") and the habit body, the sedimentations of past activities that take on a general, anonymous, and autonomous character. [..] it has affective experiences that are not merely representations; and its kinesthetic sense of its own movements is given directly.
>
> This kinesthetic awareness is made possible by a pre-conscious system of bodily movements and spatial equivalences that Merleau-Ponty terms the "body schema". In contrast with the "positional spatiality" of things, the body has a "situational spatiality" that is oriented toward actual or possible tasks. The body's existence as "being-toward-the-world", as a projection toward lived goals, is therefore expressed through its spatiality, which forms the background against which objective space is constituted. [..]
>
> The body's relationship with space is therefore intentional, although as an "I can" rather than an "I think"; bodily space is a multi-layered manner of relating to things, so that the body is not "in" space but lives or inhabits it. (Toadvine, 2019)

Merleau-Ponty's account requires substantial consideration when we talk about embodiment in videogames, since his terminology and framework make it easier to express what we are trying to affirm in this report. When we talk about embodiment, we are using Merleau-Ponty's framework, along with anecdotes and inspirations from other frameworks which we will mention.

Let's consider one of the most important pillars of this account: our existence in the world is intentional, and our body, with all of its habits and its capabilities, shapes our intentional stance towards the world, since it is our body that limits our "I can" from an endless list of possibilities down to the way we live right now. Cognition need not be thought of as perceiving, thinking (or processing), and then acting, but rather, we live in direct interaction with the world, and perception and thinking and acting are no longer separated, no longer representational, but through our long-formed habits, our spatial presence and a body schema that shapes our capabilities towards the world around us, the world appears to us directly with meanings and values.

The body schema and our ability to morph this body schema through our interactions with tools and in different contexts is vital to our discourse. Merleau-Ponty's account allows for our body schema, which is what shapes our intentional stance towards the world, to be changed as we incorporate tools and certain environments into our lives. His famous example of a blind man's stick is worth mentioning:

> "When the cane becomes a familiar instrument, the world of tactile objects expands, it no longer begins at the skin of the hand, but at the tip of the cane.
>
> [..] the cane is no longer an object that the blind man would perceive, it has become an instrument with which he perceives. It is an appendage of the body, or an extension of the bodily synthesis." (Merleau-Ponty and Smith, 1962)

Andy Clark gives a similar account when talking about our embodied experience of using virtual reality headsets:

> The infant, like the VR-exploring adult, must learn how to use internally unresponsive hands, arms and legs to obtain its goals.
>
> [..]
>
> With time and practice, enough bodily fluency is achieved to make the wider world itself directly available as a kind of unmediated arena for embodied action. At this point, the extrabodily world becomes poised to present itself to the user not just as a problem space (though it is clearly that) but as a problem-solving resource. For the world, specially when encountered via inhabited interaction, is a place in which we can act fluently in ways that simplify or transform the problems that we want to solve. At such moments, the body has become "transparent equipment": equipment that is

> not the focus of attention in use. Instead the user "sees through" the equipment to the task in hand. When you sign your name, the pen is not normally your focus. The pen in use is no more the focus of your attention than is the hand that grips it. Both are transparent equipment. (Clark et al., 2008)

And so, to summarise what we mean by embodiment as we talk about it here:

1. Cognition depends on our body as a whole, and our experiences that arise are specifically tailored by our body and its particular features.

2. Our body has an intentional stance towards the world, and this intentional stance is dependent on our habits, and is limited by the capacities of our body.

3. The "body schema" is what allows for our pre-conscious kinaesthetic awareness of our body in a "situational" sense, oriented towards possible tasks.

3 Videogaming as an Embodied Activity

This is not a simple question, and our discussion here is not to be taken as granted, of course. There are many complexities involved with attributing something as complex as "embodiment" to an activity as complex as playing videogames. This is a puzzling notion, but nevertheless, it is worth considering and thought.

Given the framework described, we can now formulate videogaming as an embodied activity. A more trivial example of what we are trying to formulate is driving cars, which is a common example used when talking about embodiment in cognitive science. When we drive a car as a proficient driver, we manoeuvre by considering what we want to do, and acting towards that direct goal without focusing on how we do this by using the gears, the clutch, the brake, pedal and the wheel, etc. We might be taking 3 to 4 actions at the same time, e.g. when reverse turning: brake in, clutch in, wheel to one side, change gears to reverse, look in the mirrors, however we are mostly thinking about where we want to go, not all the details and specifics of our interactions with the car's interface. Similar to the example of the blind man and the stick, the apparatus has become transparent and now our body schema includes the car. We decide we want to reverse and turn to one side, and given our new intentional stance towards the world that is limited and extended using the car, we consider our self to be capable of doing so. The way we question and talk about the world changes, too, we ask "do I fit here?", wondering if we can pass through a narrow passage with the car, we are now embodying a new intentional stance towards the world, and this new body schema is what gets attribution for our action.

Videogames are similar, with the difference that instead of sitting inside a car that moves spatially in the world, our human body sits in one place, but we still go places in the game-world. A proficient gamer is not concerned with the buttons they press or how they move the mouse, for example, they are directly concerned with what they do in the game-world. A new intentional stance arises

towards the game-world, that is defined by the avatar that we embody in the videogame.

Our body schema is now extended to include the avatar in the game-world, and this new body-schema limits what our human body does (just like in driving a car where some of our body is not actively used towards our goals), *we* now want to *climb* things with our new intentional body, and *shoot* the monsters and we feel real feelings of anxiety and stress (and we may even sweat) when we are playing a stealth game and we are in hiding. *We* are afraid of being found out, and when we are getting hit by enemies or falling from a height, our human body tenses, and we sometimes even get the feeling of falling dropping in our stomach (this can depend a lot on the camera of the videogame, which we will talk about). When playing a car or motorcycle racing game, our human body inevitably leans in as we are turning in the game-world. Videogames have structured worlds, with certain rules that make them predictable enough to an experienced player, and much like the real world, this can lead to us believing that we have control over the world and we can take guided actions towards certain ends. A high correspondence between our interactions with the interface that connects us to our extended body in the game-world (e.g. the game controller, or the keyboard and the mouse) and visual and proprioceptive feedback might be the key to creating a strong sense of ownership of actions. (Martin, 1995; Tsakiris and Haggard, 2005)

Besides the notion of embodiment that we have been discussing so far, there are other kinds of embodiment. Social embodiment seems to be a more ambiguous and challenging notion that must be considered with care, but consider Barsalou et al. (2003)'s account of social embodiment effects:

> "First, perceived social stimuli do not just produce cognitive states, they produce bodily states as well. Second, perceiving bodily states in others produces bodily mimicry in the self. Third, bodily states in the self produce affective states. Fourth, the compatibility of bodily states and cognitive states modulates performance effectiveness"

Real-time online videogames can exhibit similar effects, I may walk with my avatar towards a friend's avatar in the game-world and wave my hand, leading to them waving their hand, and as I start walking away, they might follow me and we may start an activity together without need for verbal or text communication, but rather only by the effect of our avatar's state of body. We have learned the affordances of our new environment and our new extended body, and that of our fellow players.

4 Camera, Avatar and Controller Relations

The camera-avatar relationship and the input interface are important factors to be considered when asking questions about embodiment of the experience, so it is necessary to consider these factors more explicitly.

(a) First-person view (b) Third-person view (c) Isometric view

Figure 1: Three common kinds of camera views

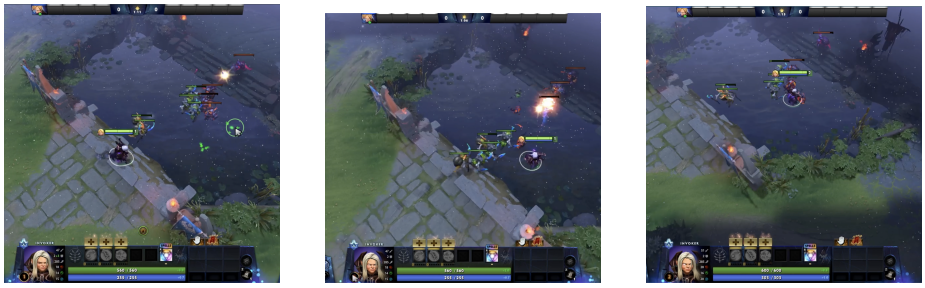

Figure 2: Camera's independence from the avatar in Dota 2

Most research around this subject seems to focus on a First-Person view, where the player is looking out through the avatar's eyes or head, only able to see the avatar's arms most of the time. This is the view adopted almost exclusively by all Virtual Reality games and many shooter games. The controller used with this type of view is either a dual-axis controller or mouse and keyboard where the character is moved with keys on the keyboard and the camera (or rather, the head of the avatar!) is moved using the mouse. (Figure 1a).

Another common view in videogames is the third-person view where the camera moves along with the avatar as the avatar moves. The camera usually has the ability to look around the avatar by rotating in its place, but never able to move away from the avatar across any axis. This view is also similarly accompanied by either dual-axis controller or keyboard and mouse where the keyboard is used to move the avatar and the camera while the mouse is used to rotate the camera. (Figure 1b)

Note that these two camera modes, albeit similar in some aspects, lend us completely different body schemas and they change our intentional stance strongly. This is best illustrated by the online multi-player videogame Dead by Daylight, where in a post-apocalyptic setting, a group of survivors are trying to survive against a killer who is trying to kill them, both of which are played by actual players. What is interesting is that the survivors and the killer use different camera views, and this is an important distinction between the two. Survivors have third-person camera which allows them to rotate the camera

and look behind them as they are running away or as they are trying to fix a broken engine to get their car running so they can run away, this also means that the survivors avatars do not move their head as the camera is moved. On the contrary, the killer has first-person camera, this means that the killer can only look in the direction that they are running in, and this allows survivors to be able to know where the killer is currently looking at, which is a significant advantage for them to be able to hide from the killer. There is a significant difference between how these two roles are played in this videogame in part because of the camera movement.

A less common, but still discussed in the literature type of camera-avatar relation is that of isometric cameras locked on the character, found in the Diablo game series. This kind of camera-avatar relation is very similar to a third-person view, with the difference that the camera is taking an isometric angle and is not controlled by the user at all, merely following the avatar. The controls used for this kind of game are usually either a dual-axis controller, or in case of keyboard and mouse, the mouse, rather than the keyboard, is used to move the avatar by issuing commands to move to a certain place. This type of movement control might seem unintuitive, however Klevjer (2012) proposes that "because the clicking happens so fast, the experience nevertheless approaches a sense of "pulling" the avatar through a tangible interface." and as such, the control interface can still create a sense of high correspondence between the player's actions and movements of the avatar, reaching a real-time synchrony as mastery of the control interface is reached.

What is common between these three camera-avatar relations is the tight coupling of the camera with the avatar: the camera always follows the avatar as the avatar moves around the world. In some cases, the camera can be rotated or moved around slightly to peek around a box while crouching for example, but almost always the camera and the avatar are in tight synchrony. Klevjer (2012) considers all of these camera modes to fall under the same umbrella of camera indirectly controlled by the movements of the avatar, as if the camera is pulled by the avatar around with an invisible string.

This group of camera-avatar relations can be considered to be intuitive and similar to how we as humans almost always have a synchrony between our vision and our body, with exception of cases like out-of-body experiences where a person sees the world and their own body from a place outside of their physical body. (Blanke et al., 2004) However, there are videogames where something analogous to an out-of-body experience happens, and these are videogames where the camera is not automatically attached to the avatar, but rather, the player has manual control of the camera. This camera-avatar relation is most characteristic of MOBA (Multiplayer online battle arena) games such as Dota 2, where the camera angle in relation to the avatar is very similar to that of Diablo, with the difference that the mouse is not only used to move the avatar, but also allows panning of the camera across the world. (Figure 2)

In these videogames you are allowed to look at the world and your avatar from any place, and given our framework, the camera is now a novel extension to our body-schema. Most embodied activities exhibit the same synchrony of

vision and body, like walking, swimming, driving a car, and in most cases of playing videogames too, however in this case, we have a new range of intentional acts available to us, through movement of the camera around the world. Our body-schema now includes a different apparatus to work with, it's as if our vision is no longer limited to our body, but rather there is a drone above us that we can see from.

This opens up the possibility of a new kind of vision interaction with the world. When the stakes are high, as is the case with e-sports, players strive for the ultimate proficiency with their new body-schema, and the result is ways of using vision that are not usual and can sometimes be cryptic for us. Camera movement of professional players tends to be very fast, and sometimes outright chaotic to an untrained eye since they want to optimise being able to scout for information while still keeping an eye on their avatar, since the avatar is still the most important part of the game, and the free-form camera movement is mostly used similar to a binocular: to scout for information.

5 Further Questions

There is a question to be asked here about how much this technical camera-avatar independence leads to actual camera-avatar independence: do players actually end up with their camera away from their avatar much, or is the camera still in synchrony with the avatar for the majority of the time, but in a manner directly controlled by the player rather than automatically.

Does the camera-avatar independence affect the embodied experience of playing this videogame, perhaps by making it more difficult to be proficient in the game, it is initially harder to extend your body-schema with this new form of vision, but what happens once you are proficient?

Competitive e-sports increase the stakes and are motivation for players to strive for utmost proficiency in a videogame, this usually leads to players being very creative and highly skilled in using the interface available to them (e.g. mouse and keyboard or a controller). In case of MOBA games with independent cameras, players reach very high action-per-minute numbers, in the last game of the largest competition for Dota 2, The International 10, the players averaged 303 actions per minute, which is about 5 actions per second, not including camera movements (camera movements are fluid and continuous and are not considered as discrete, hence their exclusion from a numerical value). (DOTABUFF, 2021)

Dolezal (2009) considers the question of action-ownership and stakes with regards to telesurgery and embodiment. She stresses the importance of a feeling of agency towards the task at hand, and proposes that high-fidelity technologies could help induce a sense of agency and ownership of action.

There is a place to ask a similar question about videogames, when the stakes are high, such as competitions with millions of dollars at stake, do players think of the actions they take in the game as their own, do they feel complete agency towards their actions in the game? What factors are at play here?

6 Conclusion

Videogames are usually formulated under information-processing cognitive models when studied in cognitive science, however on a closer look, they can be considered an embodied activity given the right framework. Here we consider Merleau-Ponty's intentional stance and body schema as a framework to formulate how a videogame might be considered an embodied activity.

Camera-avatar relations are an important factor affecting our intentional stance in a videogame, and they lend us different body schemas, from first-person and third-person camera views to an independent isometric camera that is controlled by the player. Independent cameras in videogames allow for a novel extension to our body-schema, a vision that moves freely of the rest of the body.

There are still many questions left to be explored on the topic, and rightly so, as the notion of videogames as an embodied activity is fairly perplexing and requires a lot more exploration and study until we reach a more holistic understanding of it.

References

John Robert Anderson and Jane Crawford. *Cognitive psychology and its implications*. wh freeman San Francisco, 1980.

Lawrence W Barsalou, Paula M Niedenthal, Aron K Barbey, and Jennifer A Ruppert. Social embodiment. 2003.

Olaf Blanke, Theodor Landis, Laurent Spinelli, and Margitta Seeck. Out-of-body experience and autoscopy of neurological origin. *Brain*, 127(2):243–258, 2004.

John Adam Campbell. *Video game engagement and pathology: Relationships between gaming habits and gaming experience, psychopathology, and personality variables*. The University of North Dakota, 2012.

Andy Clark et al. *Supersizing the mind: Embodiment, action, and cognitive extension*. OUP USA, 2008.

Luna Dolezal. The remote body: The phenomenology of telepresence and re-embodiment. *Human Technology: An Interdisciplinary Journal on Humans in ICT Environments*, 2009.

DOTABUFF. Match 6227492909. `https://www.dotabuff.com/matches/6227492909/farm`, 2021. Accessed: 2022-07-22.

James Paul Gee. What video games have to teach us about learning and literacy. *Computers in entertainment (CIE)*, 1(1):20–20, 2003.

NPD Group. Notable increases in both engagement and spending coming from kids. `https://www.npd.com/news/press-releases/2019/`

according-to-the-npd-group-73-percent-of-u-s-consumers-play-video-games/, 2019.

Rune Klevjer. Enter the avatar: The phenomenology of prosthetic telepresence in computer games. In *The philosophy of computer games*, pages 17–38. Springer, 2012.

Christoph Klimmt and Tilo Hartmann. Effectance, self-efficacy, and the motivation to play video games. *Playing video games: Motives, responses, and consequences*, pages 133–145, 2006.

Michael G F Martin. Bodily awareness: A sense of ownership. *The body and the self*, pages 267–289, 1995.

Maurice Merleau-Ponty and Colin Smith. *Phenomenology of perception*, volume 26. Routledge London, 1962.

Ulric Neisser. *Cognitive psychology: Classic edition*. Psychology press, 2014.

Esther Thelen. Grounded in the world: Developmental origins of the embodied mind. *Infancy*, 1(1):3–28, 2000.

Ted Toadvine. Maurice Merleau-Ponty. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Spring 2019 edition, 2019.

Manos Tsakiris and Patrick Haggard. Experimenting with the acting self. *Cognitive neuropsychology*, 22(3-4):387–407, 2005.