The Good, The Bad, and The Player: The Challenges to Moral Engagement in Single-Player Avatar-Based Video Games

Jaroslav Švelch
Charles University in Prague, Czech Republic

ABSTRACT

In this chapter, the authors create a theoretical model to analyze the challenges inherent in the implementation of moral choices in single-player avatar-based video game. Based on previous research in moral psychology and game studies, the authors investigate the relationship between the player's moral emotions and the events she causes to happen in the fictional world of a video game. The authors find that there are two factors that govern the identification with the moral content of the game's fiction: the implementation of moral agency into the game, of which they identify two basic scenarios (fixed justice and accumulation of deeds), and the style of gameplay the player chooses to follow. The authors give numerous examples from interviews, on-line discussions and gaming press of instances when players feel moral emotions about im(moral) actions they have taken in a video game.

KEYWORDS

ethics, social learning, video games, game studies, moral psychology, moral philosophy, avatar, virtual world, simulation

INTRODUCTION

"When I play a BioWare role-playing game, my characters tend to not only lean toward the nicer side, but almost immediately start twinkling with the magical pixie dust of purity. It's embarrassing, but I just make the decisions I believe I'd really make, and end up that way." - John Walker, Eurogamer.com (Walker, 2009)

"I laugh out loud when I run pedestrians over in Grand Theft Auto and get a kick out of unleashing Godzilla on my Sim City. In fact, I can't name a video-game that did evoke any sadness or true ethical dilemma in me until BioShock." - Osama, TowardsMecca.com (Osama, 2008)

In February 2009, John Walker, a well-known video game journalist, started an intriguing experiment. He revisited *Knights of The Old Republic*, a 2003 role-playing video game designed by the Canadian company Bioware and set in the *Star Wars* universe. The game lets players choose between the Light and the Dark side by carrying out good or evil actions in its fictional world. "Playing evil" results in your character having different sets of skills and looking haggard and scarred. Revisiting the now-classic game, Walker set out to become the vilest character possible—"the bastard of the old republic." His series of articles chronicle not only his bastard's deeds, but also his own reflections on the actions he took. At one

point, his character encountered a doctor who was "secretly treating the very poorest citizens, for free, against the wishes of the crimelord Davik" (Walker, 2009). Given a number of conversation options ranging from altruistic to gruesome, Walker made the bastard tell the doctor that: "If he didn't give me all the money and health packs he had, I'd report him to the authorities." Walker confesses that "it hurt to do it." "Seriously, I physically winced," he adds (ibid.).

Walker is, of course, not the first player to experience a clash of his avatar's and his own moral identity while playing a scoundrel in a video game. Neither is he the first player to feel real emotions about a fictional event in a video game. Besides a blooming academic debate on emotional engagement in video games, there are ongoing conversations on gaming forums and even a couple of "saddest moments in gaming history" lists on the Internet (TheFluffyFist, 2006).

On the other hand, as Osama from the *Towards Mecca* gaming blog reminds us, games—even the same games that make certain people sad—can be played in a calculating, ruthless, emotionless way (see Sicart, 2009). We can run over pedestrians in an organized crime simulation like the *Grand Theft Auto* series without giving the slightest thought to ethics.

In Walker's case, we can observe a certain degree of *moral engagement*, which is lacking in Osama's account. What makes the player feel or not feel morally engaged? What drives our moral choices in games and how do we relate to them? These are some of the questions this chapter is trying to answer. To do so, I develop a theoretical approach that can conceptualize moral engagement in video games and hopefully give us a better understanding of how games can provide moral experiences. I believe this type of approach is necessary to find a place for games in the project of moral education.

In this chapter, I have chosen to focus exclusively on single-player games, as opposed to multi-player, to shed more light on the relationship between the player and the designed experience of a video game, instead of analyzing interaction among players. To limit the topic even more, I focused on avatar-based games only. The avatar is "an embodied incarnation of the acting subject" (Klevjer, 2006, p.87). In referring to avatar-based games, I mean games in which the player controls one character or a small group of characters led by a main player character. The relationship between the player's and the character's morality is one of the key topics of this chapter.

My project has two complementary parts. On one hand, I will examine how game design can contribute to moral engagement. On the other hand, I will be looking at styles of gameplay that enhance or hamper one's moral engagement with a game. Most examples will be from recent and fairly recent mainstream games. The theoretical investigation will be bolstered by accounts of gamers' experiences, some of them coming from gaming blogs and discussion forums, others from interviews I conducted for the purpose of this study. Six students from my Spring 2009 game studies course at Masaryk University in Brno, Czech Republic signed up for participation in a moderated focus group. All of them are experienced players, but none had any theoretical background regarding the topic before being interviewed. The two central discussion questions that drove the interviews were: "How do you decide when you face a moral choice in a video game?" and "Do you sometimes feel bad about something you've done in a video game?" After transcribing the recording of the conversation, I translated it into English. Although it inspired some of the concepts presented here, the chapter should be considered a theoretical one, with quotes from gamers providing examples.

TO FEEL, OR NOT TO FEEL: THEORETICAL PRELIMINARIES

Moral Choice in Moral Philosophy and Psychology

Using media for moral education is nothing new. Over the centuries, literature has advanced many models for moral and immoral behavior and provided *morals* in the most literal sense. But as moral psychologist Amélie Rorty notes, literature also lacks something:

> *The role of developing the imagination—particularly through reading literature and drama—is, as may have recently argued, central to developing moral sensibilities, particularly to refining moral perception. But while such sensibility may be necessary for morality, it is not sufficient. It is as difficult to carry imagination to practice as it is to act from rules and principles. Imitation and practice are necessary for both. Unless they are expressed in the smallest nuances of practice, the principles of justice [...] can at best introduce the beneficial complexity of internal psychological conflict (Rorty, 1993, p. 38).*

Fiction obviously cannot substitute moral practice. Rorty, however, only takes non-interactive fiction into account. Making moral choices in video games is markedly different from observing moral choices in literature. As Aarseth explains in his theory of cybertext, digital games are *ergodic*—the sequence of signs that is presented in the course of the game changes according to the player's actions (1997). In other words, the player has agency to change the state of the game. In the games we are studying, this means that she has agency (and possibly, responsibility) over the avatar's actions in the fictional world of the game. At the same time, making a moral choice for a fictional character in a video game is obviously different from making a moral choice for one's self in the real world, because it remains a mediated experience.

We have, however, observed that there may be a link between someone's moral sentiment and the events in the fictional world. When this link is maintained, we can speak of *moral engagement*. To clarify the nature of this link, we must bridge the theoretical gap that lies between the two poles. The theoretical approach of this chapter is informed by two different disciplines: moral psychology and game studies. In terms of moral psychology, we will focus mostly on the developments of the last decade that contribute to our understanding of moral engagement.

Traditionally, morality was viewed as being governed by reason (Greene et al., 2001). Lately, however, new discoveries in moral psychology have caused a shift in how we approach morality (Nichols & Mallon, 2005). Neuropsychological studies, such as the one by Greene and his collaborators, have shown that immediate emotions might play a much greater role in judging (or making) a moral decision than previously thought (2001).

Greene's team's goal was to explain the paradox of the two dilemmas well-known in moral psychology. In the *trolley dilemma,* a runaway trolley is heading toward five people who will be killed unless the trolley changes course. The only way to save them is to hit a switch that will turn the trolley onto an alternate set of tracks where it will kill one person instead of five. Most people say they would hit the switch (Greene, 2001). In the *footbridge dilemma*, the trolley is, again, running to kill five people. You are standing next to a large person who is a stranger to you on a footbridge above the tracks. The only way to save the five people is to push the stranger off the bridge and onto the tracks. He will die, but his body will stop the trolley killing the others. In this case, most people choose not do so, although, on a purely rational level, there is no difference in casualties (Greene, 2001).

Greene, et al. presented test subjects with variations of these dilemmas and monitored their brain activity and reaction times using functional magnetic resonance. They found heightened activity in the parts of brain related to emotion and prolonged reaction times in the subjects dealing with dilemmas that resembled the *footbridge dilemma*, as opposed to other kinds of dilemmas. According to their findings, the kind of emotional response this dilemma produces can precede—and override—rational reasoning

(Greene, 2001).

We will get back to the footbridge dilemma, but first, let us look at the implications of this finding. The possibility of our moral decisions being governed by emotions inspired more work in the field that leaned toward a constructionist view of morality. This direction is represented by Greene himself (2002) and Prinz (2007), among others. Prinz argues that "our emotions are influencing our judgments" (2007, p. 26) and points at many examples of emotional attitudes that are impossible to justify rationally. This does not mean that there are no moral rules, but that they are constructed from emotions. For example, while thinking about committing a crime, we experience the feeling of guilt, which drives us not do so, even though it would be advantageous in utilitarian terms. This feeling is to a certain extent based on the moral development we go through in our lives. Even though rational reasoning is often involved in moral dilemmas, Greene claims that it is usually employed post hoc (2002). Overall, Greene, Prinz and other authors (Rorty, 1993; Montada, 1993) hold that our moral behavior is principally guided by moral emotions (such as guilt or dignity), rather than by fixed, objective principles.

Let us now return to the difference between the footbridge and the trolley dilemmas. Greene, et al. suggest that the footbridge situation, in which a person is pushed to his or her death, the action is more "up-close and personal," as opposed to the more impersonal action, like a hitting of a switch. They, however, conclude that the personal-impersonal distinction might only be a useful "first cut." (2001, p. 2107).

A later study by Nichols and Mallon focused on the perceived permissibility of sacrificing one actor to save five in the footbridge and trolley cases. They show that the asymmetry between the two is observed even when human actors are substituted by inanimate objects, such as wine glasses (2005). The "personal" distinction, we could argue, can be replaced with a more general factor of *immediate agency*—destroying a wine glass still gets us into an immediate contact with the destroyed object.

Nichols and Mallon also note that permissibility also depends on the magnitude of the outcome of the morally questionable actions—if billions of people are saved in the footbridge case rather than five, the choice is perceived as more permissible (Nichols & Mallon, 2005). They conclude that the assessments of impermissibility "implicate three factors: cost/benefit analysis, checking for rule violations, and emotional activations" (Nichols & Mallon, 2005, p. 539). All three will come into consideration while analyzing video games. In the next section, we will turn to game studies and look for instruments which will help us understand moral engagement in video games.

Game Studies and Moral Engagement

The second source of our theoretical framework about moral engagement in games is game studies. First, let us note that there have been previous investigations of games and ethics, including works by Pohl (2008) and Sicart (2009a, 2009b). Pohl uses methods of literary theory to outline ethical interpretation of gameplay (2008). In both his paper (2009a) and the book-sized monograph (2009b), Sicart provides thought-provoking in-depth analyses of the relationship between ethics and game design. He touches on many of the topics I discuss in this chapter, but stays mostly within the realms of rational analysis. He focuses on games as designed ethical systems and, for most part, abstracts from the imperfections and distortions that occur on the interface between the game and the player, which is the part that I am investigating in this chapter. Therefore, my view can be deemed complementary to his. I will refer to both his and Pohl's works as we move along. I will, however, try to build my argument from the moral emotion perspective instead.

First, we need to make clear a few things about video games. We have already said that digital games are

ergodic (Aarseth, 1997). By taking actions in the game, we affect the simulated world. For the purposes of this chapter, we will also adopt the view that a video game is an interplay of rules and fiction (Juul, 2006). According to Juul, the rules of the game are *real* in that they have real consequences (losing a game is a real-life fact). Fiction, which contains the representation of the fictional world in writing, graphics and sound, is not.

A relationship based purely on a game's rules is not enough to make for a moral judgment. There is hardly anything immoral about capturing a piece in checkers—unless we focus on the morals of the real-world gameplay consequences (i.e., letting a weaker opponent win a game), which is not the subject of this chapter. Without fictional context, the player actions cannot be interpreted morally. As Pohl noted, "it makes a difference if we have to arrange blocks in an optimal position or if we have to save the princess from the jaws of a monkey" (2008, p. 101).

As I discussed earlier, according to moral psychologists, emotional engagement is one of the important factors in moral decision-making (Greene et al., 2001). At the same time, there are voices in game studies that underline the role of emotion in gameplay. In her exploration of narrative capabilities of interactive media, Murray speaks of the desirability of "emotionally authentic experiences" in immersive environments (1998). Both Murray and Rusch (2007) conclude that the emotional link between the fictional situation and the player is established through *agency*—the ability to trigger various events in the fictional world by our deliberate actions. When we can interpret such actions morally (such as in moral dilemmas) and the player is emotionally engaged, we will speak of *moral engagement*.

But the player cannot enter the fictional world herself. She enters it via an *avatar*. According to Kjevler, the "avatar mediates fictional agency" and has capabilities and restrictions that "define the possibility-space of the player's fictional agency within the game" (2006, p. 87). The avatar, constrained by the game's rules, is the site of the player's agency. This agency can make us feel moral emotions about our/our avatar's actions. We will see that the player/avatar dichotomy plays an important role in our investigation. We must also keep in mind that the avatar has a fictional identity—either created by the game designers, or co-created by the player, or both—which may include the character's moral profile. Kjevler distinguishes between a *playable character,* which is the fictional character with his or her background and personality, and an *avatar,* which is the mechanism of fictional agency. We will not use two separate terms, but our use of the term *avatar* will encompass both meanings.

SETTING OFF THE BOMB: A MODEL SCENARIO

Having all the concepts in place, we can look at a hypothetical gameplay situation. At the beginning of *Fallout 3* (Bethesda, 2008), a stranger in a bar in the town of Megaton offers a mission to the avatar in return for a reward. It consists of detonating a nuclear bomb and effectively destroying the town and killing all its citizens (for obvious reasons, the mission was left out for the Japanese version of the game). The situation is of considerable complexity and can have many unexpected consequences. But let us simplify it for now and focus only on the act of accepting or refusing the mission.

The player's decision may be motivated by various considerations. There are *gameplay* concerns, such as the desire to win the game, maximize the score or explore all the possibilities the game offers. There are also *fictional* concerns. The player may hold interest in the lives of the game's fictional characters or role-play an avatar with a certain personality traits. *Contextual* circumstances are also important: the player may behave differently when being observed than when playing alone. All these factors relate to the three factors identified by Nichols and Mallon as relevant for moral dilemmas and all of them affect the level of moral engagement.

What happens when the player decides to take action? On one hand, there is the avatar, created by the player, who may or may not express fictional moral emotion prior to and after partaking in a fictional event. On the other hand, there is the player, who may or may not experience real moral emotion prior to and after partaking in a real event. When morally engaged, the player also experiences real emotion over a fictional event. (Real emotion about a real event is of course also possible, for example anger about losing the game.) Basically, given that we only include the moral emotion of *guilt* in our analysis, we could be dealing with these hypothetical situations (not necessarily featured in the game of *Fallout 3*):

[1] The player chooses to destroy the town, with a moral emotion of guilt.
[2] The player chooses to destroy the town, without a moral emotion of guilt.
[3] The player chooses not to destroy the town, avoiding a moral emotion of guilt.
[4] The player chooses not to destroy the town, not avoiding a moral emotion of guilt.
[5] The player has no choice but to destroy the town and does so with a moral emotion of guilt.
[6] The player has no choice but to destroy the town and does so without a moral emotion of guilt.

In cases [1], [3] and [5], the player experiences a moral emotion. To induce this experience, both the game and the player must work together. In [1], the potential reward might outweigh the guilt, whereas in [3] the player probably chose to follow her moral principles. The lack of choice in [5] can still be morally engaging, as we will see in the examples of *fixed justice scenarios*.

In cases [2], [4] and [6], the player does not experience a moral emotion. Both the game and the player might be the cause. First of all, the game might have not made clear the consequences of the decision or it is just badly designed. We will examine these factors in the section about the game's share in moral engagement. In [6], the player's agency is limited and therefore he might not be engaged. Lack of emotion in [2] and [4] might be caused by the player choosing a gameplay style that does not engage in fiction, such as *meta-gaming*, which we will also discuss later.

Another relevant explanation is that the player chose to perform as an avatar who would not feel such a moral emotion—in case of *Fallout 3*, there is even a list of "suggested" role-playing personalities, such as "badass wastelander," compiled on the Planet Fallout Wiki (2009). If the player is still engaging with the fiction and has a sense of agency, a tension between two moral identities arises (case [1]), which we saw in John Walker's *bastard* example. Walker was basically role-playing a villain and had to suffer through very unpleasant moral emotions. We will address these topics in the section about the player's share.

There would be twice as many examples if we decided to divide them based on whether the avatar expresses fictional guilt. To a certain extent, we can observe this in Rockstar Games' *Grand Theft Auto IV* (2008), in which the avatar Niko Bellic, an immigrant-turned-gangster, is often presented as a person who does in fact not enjoy violence. The game however, rarely offers the player non-violent options. This tension, documented by Sicart (2009a), could be most engaging in cases [5] and [6], in which the player would have to blow the town up and confront the avatar's reaction.

The bomb scenario has provided a model of possible outcomes of a moral dilemma in a video game and we will continue referring to it. Now we can look at both sides of the story, first at the game's share and then the player's share in moral engagement.

MOVING PIXELS: THE GAME'S SHARE IN MORAL ENGAGEMENT

Not all games will elicit moral emotions. But as both the previous work on games and ethics and our examples have shown, there is a growing awareness of moral issues in games. This does not necessarily mean that these games are being made with the intention of promoting moral development. Inclusion of

moral choices and dilemmas generally makes for a deeper narrative, to say nothing of the thrill of "being evil." Having a choice often implies branching narratives that cannot be revealed in their entirety on a single playthrough. This adds to replay value and sparks discussion on games' forums, where players discuss the rationale and the implications of their choices. Some of these conversations can be very engaging, as we will see later in this chapter. Even before a game is released, fans may discuss which character – good, or evil – they will choose to play, which in turn creates the much needed buzz. Especially in the role-playing genre, advertising moral gameplay has become something of a marketing instrument, not unlike graphics engines in first-person shooters. As Peter Molyneux said in regard to *Fable II* (Lionhead, 2008), a game that boasted many moral choices: "[At] the end of the day I believe choice and freedom will make you remember the experience, especially if you make a choice and there is real consequence to that choice. I think that is far more engaging than just following a linear story" (Nieborg, 2008). We should keep in mind that the fact that games marketed as having moral gameplay will not necessarily elicit moral engagement – and vice versa.

Looking for ways in which moral issues are implemented into video games, we must not forget about the quality of their fictions—writing, graphics and sound contribute to immersion by building make-believe worlds to which a player can emotionally relate. Ward, in her critique of *Mass Effect* (Bioware, 2007), argues that realistic facial expressions are a key contributing factor to the immediacy of a gameplay experience (Ward, 2008). Although we might not share Ward's opinion, we cannot dismiss the quality of presentation altogether.

As Rusch reminds us in her work on emotional design, emotional design in video games should work not only on the level of presentation, but also at the gameplay level (2007). However, let us begin by examining games that do not actually conceptualize moral choices on the gameplay level. They too can be morally engaging.

Fixed Justice Scenario

To fulfill the objectives of the game, the player often has to take moral or immoral actions within the fictional world of the game and has no choice but to do so. In that case, we can talk about the *fixed justice* scenario. This corresponds to the cases [5] and [6] of our bomb example.

In the simplest cases, the fiction of the game presents a good protagonist battling against the evil antagonist(s). Such narrative structure can justify destructive and violent gameplay in many classic examples like *Doom* (iD Software, 1993). If you think that shooting monsters is immoral, the only thing you can do is to stop playing the game. Such games contain an implicit system of morality, according to which it is moral to bring down anything that is threatening your progress in the game. Later, we will show that this morality "from above," which is rarely questioned, can be used by game designers in a subversive way.

While it might seem that no moral engagement is involved in cases where one's progression in the game is fixed, that would be a gross oversimplification. In moral philosophy, a similar question was examined by Frankfurt (1986). He concludes that the principle whereby "a person is morally responsible for what he has done only if he could have done otherwise" is flawed, as it does not account for the person's intentions prior to the action (1986, p. 143). In these games, rather than making choices, the player has to "live the values of the system and reflect upon its consequences and meanings" (Sicart, 2009b, p. 163). Sicart calls these "closed ethical games" (2009, p. 215).

Playing with the player's intentions, linear games will use various deception techniques that make the player feel moral agency, although the rules do not offer moral choice. The first one is the illusion of

choice, which is often just a method of making the game seem bigger that it is. It may, however, be a very compelling way of representing the futility of one's efforts. For example, in *Silent Hill 2* (Konami, 2001), the player experiences the illusion that he can save a fictional character who is being attacked by the game's antagonist, the violent Pyramid Head. Although the avatar cannot accomplish this, the player can make the avatar (James) try. Such futility, in real life just as in video games, easily provokes the feeling of guilt. As Rusch puts it in her critique of the game:

> *As a player, I felt guilty for saving my own ass and leaving Maria to die. I tried to fight Pyramid Head, but he was far too strong for me and I got killed every time I tried to stand up to him. Running, behaving like a coward, was my only hope [...].* (Rusch, 2009)

The other technique can be tentatively called the blind follower model. In games that lack narrative choices, the player tends to control the avatar based on what he is told by "the game" (be it in instructions or in-game briefings). She accepts an external justification for the actions she is making the avatar take and, in a way, delegates his or her moral responsibility to a higher authority (see Montada, 1993). As a result, a narrative structure emerged in video game storytelling which revealed this authority to be immoral and manipulative (Weise, 2009).

Such narrative strategy can be used to great effect. The narrative first-person shooter game *BioShock* (2K Games, 2007) has a limited moral choice (which will be discussed further), but mostly consists of a linear narrative that takes place in the ruined underwater city of Rapture. In the first portion of the game, the player is contacted via radio by a person named Atlas who insists he is an ally, and that he needs help. There is no way of knowing whether he is trustworthy, but the game constrains the player to do only what he is told. Halfway through the game, Atlas turns out to be the game's villain. The avatar was nothing more than his puppet and the game deceived the player into adopting the ethics of the villain. Such revelation can be a powerful generator of moral emotions.

A similar scenario takes place in the acclaimed *Shadow of the Colossus* (Sony, 2005). In this game, the avatar Wander is promised that after slaying 16 huge creatures—colossi—his dead love will come back to life. These colossi, however, are not portrayed as "evil", but as peaceful, solitary forces of nature. As the player leads Wander to kill the colossi, his appearance deteriorates—he grows paler, his hair gets darker—which might be interpreted as externalization of the avatar's guilt. However, carrying on his nightmarish task is the only thing he—and the player—can do. He is even rewarded by becoming more powerful.

In the case of *Shadow of the Colossus*, the player is not really blind. She can see that Wander is governed by desperation rather than morality (Sicart, 2009b). He wants to kill the colossi, while the player might not. As Dave P. put it in the comments on the *Joystiq* magazine forum:

> *[I felt] sadness not only because I knew I was 1/16th closer to the end of what was a wonderful gaming experience, but sadness because I wasn't sure my character was doing the right thing. I mean, who am I to be destroying these majestic things in order to cheat death? (Dave P., 2005)*

The fiction of the game is powerful enough to elicit moral engagement. That does not mean that a realization such as Dave P.'s comes automatically—it does require an open mind. At least, it comes up very rarely on the game's official Sony forum.

All things considered, we might conclude that in the blind follower model, a tension arises from the possible clash between the player's moral values and the actions she is forced to take. This tension can be frustrating, but it can also be emotionally powerful. We have seen that games using both this model and the illusion of choice technique may provide the player with profound moral experiences, although they might not be considered as having moral gameplay in the traditional sense of offering choices. .

Accumulation of Deeds Scenario

In the simulated fictional worlds of video games, the choices we make for an avatar have consequences. Usually, games that advertise moral gameplay will let the player make moral choices and use rudimentary morality systems to give the player feedback on the moral value of (some of) her actions. Games such as this will confront the player with a series of situations corresponding to cases [1] to [4] in our bomb example.

The adventure game *I Have No Mouth And I Must Scream* (Cyberdreams, 1995), the first game to be marketed as addressing complex moral issues, introduced the “spiritual barometer”. Based on the player's actions, the multiple avatars in the game were evaluated from a moral standpoint by placing them on a scale of the barometer. Since then, many games have used a single sliding scale of good and evil. The position of the player on such scales may determine non-playable characters’ attitude toward the avatar(s) or trigger other events in the game world. Often, a certain branch of narrative opens based on a moral or other choice.

I will use the science fiction role-playing game *Mass Effect* as one of the examples of what is challenging about designing moral gameplay and what can go wrong. It is an ambitious game that is, simultaneously, indebted to the traditions of role-playing, narrative branching and morality scales. Combining psychological and moral profiles, *Mass Effect* gives the player points on the Paragon (meaning adheres to the law) and Renegade (meaning neutral to the law) scales. Both are considered “good” though, as the Paragon is someone who follows the rule of law and acts with tact, whereas the Renegade is someone who does things outside of the rules, or is not always as diplomatic. The assignment called The Negotiation, for example, unfolds in the following way:

> *Admiral Hackett needs Commander Shepard [the avatar] to negotiate with a warlord named Lord Darius. When you arrive at the base Darius will begin the negotiations on the offensive [...]. As you proceed through the negotiations, you can choose to appease Darius and end the mission peacefully. Alternatively, at any point you can decide that you’ve had enough of “Lord” Darius and simply kill everyone. You receive 8 Paragon points for ending the conversation peacefully, or 25 Renegade points for killing them all.* (Mass Effect Wiki, 2009)

Although systems like this overtly offer moral gameplay, they may suffer from drawbacks that threaten the link of moral engagement. We will divide them into four categories, arguing that moral gameplay may be: imposed, quantified, polarized, predictable and inconsistent.

First of all, the moral values of particular actions are imposed by the game designer. Why 8 points for a peaceful solution and 25 for killing everybody? This sort of arbitrariness demasks the system and reveals that the consequences come not from the fictional world, but from the designers.

Secondly, the feedback is usually quantified and the consequences (or at least some of them) arrive instantly. Not only is it a crude oversimplification, but it also becomes subject to manipulation to achieve gameplay goals. To access some assignments in the game, for example, you have to have a certain number of Paragon and Renegade points. Thus, gameplay motivation and moral motivation can get mixed up—Sicart notes that gameplay of this kind is not ethical, but “merely statistical” (2009b: 209).

Thirdly, the moral gameplay is too polarized. The variety of choices might not accommodate the moral profile that the player decides the avatar must follow. If the choices are just polar opposites of good and evil, they might not account for the player’s intention (why not just kill Darius and let others live?). This

can threaten moral engagement.

The fourth point of criticism is that moral choices are predictable. In *Mass Effect*, like many role-playing games influenced by the output of BioWare over the late '90s, the moments of choice occur almost exclusively in conversation with non-player characters. But emotional design can hardly be successful if the game screams that "a moral choice is coming." Speaking of *Mass Effect* (in an article unrelated to the *Bastard*), John Walker says:

> *I'm not sure, but I don't know if RPGs are quite the place to try and resolve the most controversial and divisive subjects of our day. Well no, that's not true at all. They could be, but perhaps in a setting slightly more dignified than as a result of overhearing a conversation on a street, and then immediately being given life or death decisions to make for complete strangers.* (Walker, 2008)

From the perspective of moral philosophy, the incredibly tough decisions such as Sophie's choice (mimicked by a similar choice between two crew members in Mass Effect) are indeed borderline cases. As Greene puts it in his moral philosophy dissertation:

> *Naturally, moral dilemmas of this kind receive much attention, but we shouldn't lose sight of the fact that most moral judgments—the ones we don't bother to talk about, the ones that are just "common sense"—are not hard at all.* (Greene, 2002, p. 172)

We should not be surprised that it is these tough decisions that are often implemented in games. They require a considerable amount of reasoning and weighing of consequences—a thought process which is ultimately game-like.

But moral gameplay is often relegated to small, isolated islands. As one of the interviewed gamers, Jarek, put it: "I tend to play differently and experiment a lot, when I know the game is not watching." Often, the game can wink a lot and then gaze with great fervor, especially at the points where its narrative is branching. Then, the player's moral engagement is fleeting and erratic.

The fifth point, inconsistency, is related to the previous one. As each option in a choice-driven game has to be specifically written and designed, integration of moral questions into gameplay and fiction is a challenging task. It might break the consistency of gameplay, fiction, and ultimately, overall poetics. In *Mass Effect*, for example, you are occasionally faced with moral choices, but you *cannot* shoot civilians (it has no effect whatsoever) and most enemies are nothing more than cannon fodder, as they don't get the special Darius treatment. There is no moral consequence to their deaths beyond completing a mission.

Let us look at another example. In *BioShock*, the player encounters the mutant Little Sisters, girls bred to be living containers of ADAM, a substance which is essential for gameplay as it allows the avatar to become more powerful. When you encounter a Little Sister, you can choose to either "harvest" (and kill) the Little Sister or "save" her and let her return to her community. Both options have their benefits, but the eventual outcome is rather balanced. In terms of actual gameplay, there is very little difference whether you decide to save them. Nevertheless, while harvesting Little Sisters counts toward the accumulation of your (im)moral deeds, the fact that you are killing hordes of other mutants along the way is not questioned in any way.

We might dismiss all these criticisms as elements of genre, and generally, game design conventions. The players are accustomed to them and, as a result, ignore them, or use them as guideposts for entering into a new experience. Also, the games that suffer from them are often more thought-provoking than games that do not attempt moral gameplay at all. But breaking these conventions of division of "normal" and "moral" actions can bring up surprising moments that make us realize our moral engagement once again.

The role-playing game *The Witcher* (CD Projekt, 2007), for example, builds on moral ambiguity. Choices come at predictable moments, but the game usually lets the player choose between two camps neither of which is morally sound: for example, you can either join the despotic and bureaucratic Order, or the reckless and violent Elven guerrillas. This makes the player question the avatar's (and her own) values. Moreover, the consequences of the player's decisions take place long after the action is taken, making it impossible for player to calculate with the outcome.

In another example, the otherwise linear stealth action game *Metal Gear Solid 3* (Konami, 2004) fiddles with the concept of the game as a *hidden moral spectator* in a very inventive way: In a haunting scene toward the end of the game, it shows the player the ghosts of all the people (mostly soldiers) the avatar, Snake, has killed throughout the game:

> *Trapped in the twilight between life and death the player is visited by visions of everyone they have killed throughout the game. The player cannot hurt any of the ghosts, but the ghosts can hurt the player. Snake must simply endure all the pain and agony he has caused, facing each and every enemy soldier he has killed, all bearing the scars of exactly the way the player chose to kill them.* (Weise, 2004)

By delaying the confrontation with the consequences of the player's actions and making them a little more unpredictable, the relative rigidness of scale-based morality systems can be prevented. How else can moral gameplay be enhanced?

Let us start with another concept taken from moral philosophy. According to one of the morality scholars we have chosen to follow in this chapter, morality is intersubjective: "the meaning of obligation is socially defined" (Tugendhat, 1993). Therefore, a game that intends to have a robust and consistent morality system should strive to simulate the social dynamics that define it and create a world that reacts to our moral agency. Such simulation could bridge the local and global consequences by introducing non-player characters that are able to react to all morally relevant avatar's actions individually and at the same time also pass their judgment on to other non-player characters. This emergent morality system would deal with the predictability and inconsistency problems. Games already exist that try to introduce a "social dynamics engine" (Koo, 2009), such as *Fable* (Lionhead, 2004) and *Fable 2* (Lionhead, 2008). The implementation of such an engine is however very limited, and the game presents the player with inconsistent reactions from the non-player characters (NPCs), the inhabitants of the fictional world not controlled by the player. (Nieborg, 2008). Before concluding the section about the game's share, let us review the possible strands of moral gameplay design in the following chart.

Scenario	**Techniques**
Fixed justice	*illusion of choice*
	blind follower
Accumulation of deeds	*morality scales*
	branching narrative
	hidden spectator
Emergent morality system	*social dynamics engine*

Table 1. Scenarios of moral gameplay and game design techniques they can utilize

DON'T WATCH ME STEALING CARS: THE PLAYER'S SHARE

No matter how well a gameplay experience is designed, it takes both the game and the player to establish an emotional link of moral engagement between her actions and the game world. As I mentioned in the introduction, the player may choose her style of play based on gameplay, fictional and contextual concerns. The gameplay-oriented styles include strategizing, meta-gaming and exploration and fiction-

oriented styles can be divided into honest play and role-playing. I will examine contextual factors separately. As all of the styles mentioned above affect moral engagement, let us now deal with them one by one, while keeping in mind that a player can switch from one style to another in the course of gameplay.

Strategizing is a predominant orientation on the game's rules, while ignoring the fiction. Earlier, we have adopted the view that the game is an interplay of fiction and rules, which are "real" in the sense that they determine the outcome of a game. As the "real" motivation of the player is usually to win the game, she might make her choice according to what style of gameplay suits her better, not according to moral concerns.

This kind of strategizing is made easier when the game's morality scales are demasked. Although Tugendhat, for instance, posits the moral ability outside of other abilities (1993), in games, morality is often just another "stat" (Sicart 2009b). We touched upon this briefly in our discussion of *Mass Effect* and we can expand the argument using another example.

In the classic role-playing game *Baldur's Gate II* (Bioware, 2000), the player is informed about the change in her reputation right after having made a moral choice. The reputation is represented numerically, the higher the figure, the more virtuous the avatar is. This allows for an instrumental use of the morality system—knowing that a certain figure on the reputation scale will change some NPC's attitude toward the avatar, the player may decide to act in order to achieve that figure, retro-fitting her decisions to her goal.

For example, one of the interviewed gamers, Jan, killed innocents in Baldur's Gate II just to lower his reputation enough so that he could hire strong non-player characters who are evil and would not join him otherwise. Otherwise, he would not usually choose the "evil" options when given a choice. This mixing up of gameplay and narrative decisions might be, to a certain extent, prevented by the game offering balanced outcome of all possible options: even in *BioShock,* your decision not to harvest the Little Sisters is eventually compensated by gifts.

The term *meta-gaming* has been established in gaming discourse as finding one's own goals in the game, which the designers did not originally intend—or playing with the system. If we are, for example, competing in how many pedestrians we run over in the motorized crime simulation game *Grand Theft Auto IV* (Rockstar, 2008), we might want to give up on any responsibility in the fictional world, if only for a moment. The player might not even consider such moments "proper gameplay" and when retelling the story of her progression through the game, she will be likely to leave this part out. Meta-gaming is like stepping out of the avatar (at least his fictional context, the playable character), while still controlling it. The disconnect naturally means that the player is not morally engaged.

In his classic typology of players, Bartle mentions the "explorer" type (1996). When compelled to discover all the content in the game, players often employ the gameplay style of exploration, not only in terms of discovering the limits of the game world, but also in terms of finding out all possible outcomes of different choices. The players save and reload games after a major decision and choose an option depending on its consequences. If the outcome is apparent immediately after the decision, the player can often reverse it by re-loading the game, potentially disengaging herself from the fiction in the process. *The Witcher*, for example, prevents this by delaying some of the consequences by several hours of gameplay.

In terms of fiction-related styles of gameplay, we can distinguish between two basic ones: honest moral gameplay and role-playing. Honest gameplay is the most straightforward kind of moral engagement. We

project our moral code into a fictional world, reaffirming our moral beliefs based on our moral emotions. For those of us who do not wish to wreak havoc in post-apocalyptic towns, this would correspond to the case [3] of our bomb example.

Examples abound. In *Bastard of the Old Republic*, Walker admits that he tends to play as if he was in the avatar's shoes. One of the interviewed students, Jan, also said: "When I played [the game] for the first time, I played as if it was me." Still, he acknowledges that it is just one of the possible styles of gameplay—a first time one. Later, the time might come for role-playing.

When role-playing, we take on a moral profile different from our own. As Jarek said, "you can try out what you would not normally do." And this is, obviously, one of the biggest attractions of video games. Role-playing a character with a different moral profile can be both fun and an interesting experiment. As a player who calls himself "nfs nole" puts it on the official *Mass Effect* forum: "I have more fun playing a ruthless character." This would fit in with the case [2] from our bomb scenario.

The motives for role-playing and the relationships between the player and the avatar can be very intricate and idiosyncratic. In the rhetoric of the posters on the *Mass Effect* forum and the in-game YouTube videos comments, there is an obvious distinction between the player and his avatar ("my Shepard," Shepard being the last name of the avatar, whose gender can be selected at the start of the game)—especially when the poster has finished the game repeatedly with different, customizable avatars:

> *Definitely renegade is fun, but for this playthrough I wanted to make it seem like a realistic movie with Shepard being totally ruthless only when he needs to be, when the situation requires it.* (Stealthspectre, 2008)

As we can see, as long as the game allows for it, the role-playing style of gameplay can be very nuanced. Still, many players encounter difficulties when trying to role-play a ruthless character, because the disconnect between the player's and the avatar's morals does not necessarily mean loss of moral engagement. Effective visual and emotional design may re-connect the player with the character he is role-playing. That is what happened to John Walker while trying to be a bastard. The same thing, represented by the case [1] from our bomb example, happened to Osama from TowardsMecca.com, while playing *BioShock*:

> *Almost immediately afterwards [the first playthrough], I started the game over with the intention of beating it once more but by harvesting every little sister this time. I confess I did do it the first time, but could not continue after that. The action of harvesting them made me feel ridiculously guilty.* (Osama, 2008)

We might concede that what makes *BioShock* so emotionally resonant is first and foremost the imagery. When choosing to harvest a Little Sister, the avatar takes the writhing little child into his hands and effective devours her, extracting from her the sea slug which had been implanted into her body and which contains the precious ADAM. No matter how mutated the sisters are, the visual representation of the act is very vivid and disturbing and reminds us of the *immediate agency* of the footbridge dilemma.

Another example is the "I Remember Me" subplot in *Mass Effect.* In this assignment, commander Shepard is asked to calm down a young woman, traumatized after years as a slave to an alien race. She is threatening to kill herself and Shepard is the only one who can talk her down. Unless Shepard proceeds in a very sensitive way, the woman will shoot herself. Although available only to certain players (based on some customization choices prior to playing the game) and being rather insubstantial in terms of the master narrative of the game, it has gained a lot of attention on the forums. As *dreyshock* confides:

> *I don't think I could bring myself to play renegade for this particular sub-plot. It'd be too heart-wrenching to see her commit suicide or attack Shepard, which is what I suspect would happen.*

(Dreyshock, 2008)

For another player, it dramatically changed the concept of his or her avatar, as if a new aspect of Shepard's personality had been revealed:

> *"I Remember Me" is probably my personal favorite assignment in the game. I got to see a completely different side of my Shepard. She was so harsh up until that scene, and changed my mind on a lot of my decisions.* (Usasoldiern, 2008) [typos corrected by J.Š.]

In this example of the tension between the avatar and the player, the player basically role-played character development. Thanks to moral engagement on the side of the player, Shepard, who is an elite soldier and is both voiced and animated accordingly, found his or her soft side.

Both the devouring of the Little Sisters and the "I Remember Me" sequence are, of course, carefully staged examples. But even the more mundane gameplay actions can be interpreted morally. Although we often abstract from moral considerations while playing games such as *Grand Theft Auto* (especially when it is so difficult to not run people over), an observer can make us change our mind. Probably the most interesting contextual factor of moral engagement, spectators can reinforce the player's connection with the fiction of the game. As our interviewee Jarek said:

> *I was playing* Grand Theft Auto *and my son came around and asked me: What are you doing, daddy? Are you stealing cars? Then I realized I should not be doing it, at least not in front of him.*

Spectators might not be participating in the gameplay, but they can easily engage with the game's fiction. Another interviewee, Štěpán, noted:

> *I make different decisions when I'm controlled by something out of the game—in this case, it might be my girlfriend who identifies with the avatar and tends to take the in-game events more personally. This makes me more "moral" in the game.*

I believe that this *spectator effect* deserves deeper investigation. This effect is also in place in multi-player games, but those bring about numerous other concerns that we cannot deal with in this chapter.

Let us conclude this section by saying that players choose to morally engage or disengage based on what the game provides and their own goals. This does not mean that they are immune to the fiction or to moral concerns. We can see the different gameplay styles and their relation to moral engagement in the following chart.

Player concern	**Gameplay style**	**Moral engagement**
Gameplay	*strategizing*	-
	meta-gaming	-
	exploration	-/+
Fiction	*honest play*	+
	role-playing	+/-
Context	*spectator driven*	+

Table 2. Gameplay styles related to moral engagement, categorized by their driving player concern. Presence of moral engagement is indicated, but these statements should not be considered absolute, as the player can always be caught off guard.

CONCLUSION

The specific experience of moral engagement in video games appears at the interface between the game and the player. We have observed different design strategies and styles of gameplay that affect moral engagement. These are not entirely independent: the fixed justice scenario, for instance, rarely allows for role-playing and the accumulation of deeds scenario, to a certain extent, prevents meta-gaming.

I have also identified a number of challenges to moral engagement in video games and made it clear that

moral engagement and emotional experience is not guaranteed simply by solid game design. We have seen that it is foolish to demand moral responsibility from the player for everything immoral she does in the game, but we have also shown that clever design can make the player reconsider and contemplate her choices. There is rich evidence in interviews, forums and the gaming press that players can indeed be moved and that they will reflect upon the deeds of their avatars.

We cannot automatically consider video games as superb tools in moral development. But they can frame simple moral thought experiments (like the trolley and footbridge dilemmas) with a wider fictional context and provide a unique experience that can reveal to players through their own moral emotions how they relate to different moral codes. Although it is not easy to integrate moral issues into a game, setting them in a familiar fictional world adds a vividness and immediacy unprecedented by other media.

How can we expand on the potential of the medium? First, I believe that game scholars should conduct further research of actual player experience. We should not take a game's moral design as a given and, instead, study particular ways of engagement by the methods of player ethnography and discourse analysis of players' discussions and write-ups. This way, game designers can see how their respective approaches to moral gameplay resonate in players. It can make them reconsider the long-standing conventions and clichés that are used in design of moral gameplay and bring forth new and gripping emotional experiences. In the end, even playing evil in a video game can make one a better person.

REFERENCES

Bartle, J. (1996). *Hearts, clubs, diamonds, spades. The players who suit MUDs.* Retrieved January 10, 2009, from http://www.mud.co.uk/richard/hcds.htm

Dave P. (2005). *Joystiq review: Shadow of the colossus, discussion comment.* Retrieved July 16, 2009, from http://www.joystiq.com/2005/11/13/joystiq-review-shadow-of-the-colossus/

Dreyshock. (2008). *Mass effect (PC) insanity vanguard P37: I remember me, discussion comment*. Retrieved March 20, 2009, from http://www.youtube.com/watch?v=YzKYHi12pZ8

TheFluffyFist. (2006). *Saddest moments in video games - Part I.* Retrieved August 28, 2009, from http://www.theaveragegamer.com/2006/11/15/saddest-moments-in-video-games-part-i/

Frankfurt, H. (1986). *Alternate possibilities and moral responsibility*. In J.M. Fischer (Ed.), *Moral Responsibility* (pp. 143-152). Ithaca, NY: Cornell University Press.

Greene, J. D., Sommerville, R. B., Nystrom, L. E., Darley, J. M., & Cohen, J. D. (2001). An fMRI investigation of emotional engagement in moral judgment. *Science*, 293, 2105-2108.

Greene, J. D. (2002). *The terrible, horrible, no good, very bad truth about morality and what to do about it.* Unpublished doctoral thesis, Princeton University, Princeton.

Juul, J. (2006). *Half-real.* Cambridge, MA: MIT Press.

Klevjer, R. (2008). *What is the avatar.* Unpublished doctoral thesis, University of Bergen, Bergen.

Koo, G. (2008). *Talk on games, morals and ethics.* Presented at MIT. Retrieved January 10, 2009, from http://blogs.law.harvard.edu/games/2008/11/18/talk-on-games-morals-and-ethics/

Montada, L. (1993). *Understanding oughts by assessing moral reasoning or moral emotions*. In G. Noam & T. Wren (Eds.), *The Moral Self* (pp. 292-310). Cambridge, MA: MIT Press.

Murray, J. (1998). *Hamlet on the holodeck*. Cambridge, MA: MIT Press.

Nichols, S., & Mallon, R. (2005). Moral dilemmas and moral rules. *Cognition*, *100*(3), 530-542.

Nieborg, D. (2008). *Morality and gamer guilt in fable 2.* Retrieved January 10, 2009, from http://blogs.law.harvard.edu/games/2008/11/19/morality-and-gamer-guilt-in-fable-2/

Osama. (2008). *Bioshock: The choice that mattered*. Retrieved February 15, 2009, from http://towardsmecca.com/2008/08/28/bioshock-the-choice-that-matters/

Planet Fallout Wiki. (2009). *Fallout 3 Role-playing.* Retrieved March 10, 2009, from http://planetfallout.gamespy.com/wiki/Fallout_3_Roleplaying

Pohl, K. (2008). Ethical reflection and involvement in computer games. In S. Günzel, M. Liebe, & D. Mersch (Eds.), *Conference Proceedings of the Philosophy of Computer Games 2008* (pp. 92-107). Potsdam, Germany: Potsdam University Press.

Prinz, J. (2007). *The emotional construction of morals*. Oxford, UK: Oxford University Press.

Rauch, P. (2007). *Playing with good and evil. Video games and moral philosophy*. Unpublished Master's Thesis, MIT, Cambridge.

Rorty, E. (1993). *What it takes to be good*. In G. Noam & T. Wren (Eds.), *The Moral Self* (pp. 28-55). Cambridge, MA: MIT Press.

Rusch, D. (2007). *Case study: Emotional design of the videogame silent hill - restless dreams*. Presented at MIT5 Conference. Retrieved January 10, 2009, from http://web.mit.edu/comm-forum/mit5/papers/Rusch.pdf

Sicart, M. (2009a, forthcoming). The banality of simulated evil. *Ethics and Information Technology.*

Sicart, M. (2009b). *The ethics of computer games.* Cambridge, MA: MIT Press.

Stealthspectre. (2008). Mass effect (PC) insanity vanguard P37: I remember me, discussion comment. Retrieved March 20, 2009, from http://www.youtube.com/watch?v=YzKYHi12pZ8

Usasoldiern. (2008). I remember me. Mass effect community forum post. Retrieved March 15, 2009, from http://masseffect.bioware.com/forums/viewtopic.html?topic=613268&forum=123&sp=15

Walker, J. (2008). *Morality tales: BioWare versus the issues*. Retrieved February 15, 2009, from http://www.rockpapershotgun.com/2008/08/26/morality-tales-bioware-versus-the-issues/

Walker, J. (2009). *Bastard of the old republic.* Retrieved February 22, 2009, from http://www.eurogamer.net/articles/bastard-of-the-old-republic-article

Weise, M. (2004). *Metal gear solid 3: Snake eater.* Retrieved February 22, 2009, from http://www.gamecritics.com/review/mgs3/main.php

Weise, M. (2009). *Totalitarian by design.* Unpublished manuscript.

Games

2K Boston (2007). *BioShock*. (Windows).

Bethesda Softworks (2008). *Fallout 3*. (Windows).

Bioware (2000). *Baldur's Gate II: Shadows of Amn.* (Windows).

BioWare (2003). *Star Wars: Knights Of The Old Republic*. (Windows).

BioWare (2008). *Mass Effect*. (Windows).

CD Projekt (2007). *The Witcher*. (Windows).

Cyberdreams (1995). *I Have No Mouth And I Must Scream*. (MS-DOS).

iD Software (1993). *Doom.* (MS-DOS).

Konami (2001). *Silent Hill 2*. (Playstation 2).

Konami (2004). *Metal Gear Solid 3: Solid Snake.* (Playstation 2).

Lionhead (2004). *Fable.* (Xbox).

Lionhead (2008). *Fable II.* (Xbox 360).

Rockstar North (2008). *Grand Theft Auto IV.* (Rockstar, 2008).