Bachelor Degree Project

UNIVERSITY
OF SKÖVDE

MORAL MEASUREMENT STICK

Moral philosophy in games

Bachelor Degree Project in
Media Arts, Aesthetics and Narration

30 ECTS
Spring term 2022

Kylar Eriksson

Supervisor: Karin Brygger
Examiner: Lissa Holloway-Attaway

Abstract

How moral alignments are used in video games is an issue under discussion. This study looks at two methods of presenting it to the player. One is covert and tracks alignment behind the scenes, and the other shows how the score changes according to choices. 20 participants played either version of the artefact, an interactive text story, where the difference is the method for alignment tracking used. Survey responses are then used to establish the players' perceived enjoyment and engagement with the story, as well as their engagement with moral dilemmas that were a part of the artefact.

Keywords: video games, morality, narrative systems

Table of Contents

1 Introduction 1

2 Background 2

2.1 Designing moral systems 2
2.2 Engaging with player ethos 5
2.3 Psychology 7
2.4 Theology 7
2.5 Benefits of morality systems in games 9

3 Problem 11

3.1 Method 12
3.1.1 Artefact 12
3.1.2 Survey 14
3.1.3 Distribution 16
3.1.4 Additional Issues 16

4 Implementation 18

4.1 Pilot Study 18
4.2 Progression 19

5 Evaluation 25

5.1 The Study 25
5.2 Analysis 37
5.3 Conclusions 40

6 Concluding Remarks 43

6.1 Summary 43
6.2 Discussion 43
6.3 Future Work 44

References 46

7 Appendix I

7.1 Appendix A I
7.2 Appendix B VIII

1 Introduction

This study aims to compare morality systems used in video games. The morality meter, a visual representation of one's moral alignment, has been criticized for its lack of nuance and propensity to force players into certain ways to play. Another system uses covert feedback to inform the player of the consequences of their actions like the game world changing or a character's attitude towards the player avatar adapting to the new information. This system may still use morality meters but such is not shown to the player. The main difference between the two systems is that one explicitly ties the world and story to morality meters, showing this to the player, while the other aims to hide the tracking of alignment to make the world and story more versatile. They both have their strengths and weaknesses both for game design and as a narrative tool. Even if the morality meter is being phased out in many new games, it is a notable part of gaming history and does have its uses. The study will mainly base itself on the user perspective and compare how much the user enjoys the morality meter vs the more covert feedback system.

To investigate which of these systems is the most enjoyable and engaging for the player, the study uses two versions of an interactive text game written for the study. For its construction, various game and narrative design elements have been used to engage with the participant's ethos. Furthermore, moral philosophy and theological theories have also been employed to inform the moral dilemmas in the game. Participants play one version of the game. After this, the participants answer a survey with questions pertaining to their experience.

2 Background

An aspect that digital games have over other forms of media, like movies or books, is that they can engage with the player in a more personal manner. As the player of the game tends to embody an avatar in the game world, they are present within the narrative (Hodge, Taylor, & McAlenery, 2019). This makes them an actor, rather than simply a spectator of the actions of others. The film playwright can tell a story that urges the viewer to pass judgement on the characters within it or to empathise with others. The game writer asks the player to commit the actions themselves, and often to judge those actions as well after introspection. Some games ascribe morality to certain actions. These can be presented to the players in different ways. One archetypal method is a morality meter (Formosa, Ryan, Howarth, Messer & McEwan. 2022), represented by a numerical score or a visual representation like a bar that fills up. The other method is by feedback from the game world, like characters reacting or the world changing according to the player's choices (Formosa, Ryan & Staines. 2016). A game with morality meters may still have that feedback, meaning the world reacting, but it will be clear that this is tied to a meter. This study will outline different perspectives of morality, as based in societal norms and ethos (Consalvo, Busch & Jong, C. 2016), philosophical ethics (Greene, 2009), or theological ethics (Jun, 2005)(Cohen & Rozin, 2001). This may help to understand how developers can engage with player ethos in digital games. The study will also examine how games can be constructed to achieve this goal, with an emphasis on morality meters, as well as a covert exchange of moral judgement and feedback between player and developer. It will also consider some ways that games can influence a person's perception of right and wrong.

2.1 Designing moral systems

The morality meter is a mechanic in video games which aims to track and present the player's morality within the game. This usually looks like a scale that goes up or down from a neutral middle toward the two sides. The meters may also only measure one sort of morality where it fills from the bottom and goes up, or simply be stated in numerical points. Morality meters are hotly debated by both academics and players, often criticized for taking away the nuance of ethics to be replaced by a simple good vs evil dynamic (Formosa, et.al. 2022). Morality meters are also often instrumentalized to reach desired outcomes. In this way, it is not a representation of the player's morals, but a mechanic to "fill up", much like experience points or a wallet. Morality meters are also limited by the categorisation of certain actions as good or evil, which will usually be based on the designer's own views. Popularized mainly by the

developer Bioware in the games *Star Wars, Knights of The Old Republic* (2003) and *Mass Effect* (2007), the morality meter filled up in the games depending on the player's dialogue choices. These were aligned with a certain side represented in the morality meter. The system almost required the player to stick to one side, as doing so led to the best rewards. Formosa uses the term *ethical cognitive friction* to explain some players' dissatisfaction with the system. Unless their personal morality is represented in the game, they can experience moral disengagement. This is also a contributing factor to not engaging with the game's narrative, but instead aiming toward maximizing rewards that may be available by using the systems correctly. Formosa (2022) (Formosa, Ryan & Staines, 2016) considers there to be two types of players; the reactive, who optimizes the game and its systems; and the reflective, who engages with the game's narrative semantically and makes choices that make sense in the context of how they play. A positive aspect of morality meters is that it is good at giving feedback in a way that isn't too costly. The other way to give feedback is that the world reacts to the player's actions, by characters commenting on it or the world changing as a consequence of the actions taken. This can be more intensive for every part of the development team, as they may require new assets or dialogue. According to Formosa's study (2022), it would seem that most players lean towards the reflective type when engaging with video games. The costlier method is therefore likely to be preferred as this creates the impression of a dynamic world that reacts to the player's *actions* instead of their *moral score*.

Formosa also analyses the game *Papers, Please* (2013) in *Papers, please and the systemic approach to engaging ethical expertise in videogames* (Formosa, Ryan & Staines. 2016). In contrast to games like *Mass Effect*, this game leans heavily on its mechanics to tell a story. *Mass* Effect is an epic space opera about a military commander trying to unite a tumultuous galaxy against an even bigger threat. The story stretches across three long games and is backed by the money of large corporations. *Papers, Please* is an indie game set only in a small area. The player avatar is a border control agent for an authoritarian nation, and has to make sure that undesirable people do not enter the country. The game uses the bureaucratic way of seeing migrants as a passport or a paycheck to dehumanize them. The rhetoric is that anyone can become an uncaring cog. This is supported by the game's mechanics, all of which have ethical dilemmas. Formosa highlights that *Papers, Please* differs in how morality is handled compared to *Mass Effect*. In *Mass Effect*, choices are made at pivotal moments in the story. The system is more overt. *Papers, Please*, on the other side, is covert. The game is always questioning the player with small decisions.

This study will be designing an interactive story. For this, two sources have been used to understand the composition of a narrative that addresses ethical dilemmas in a nuanced way. On the blog Pluralsight (2015) five steps can be found. First and foremost, if the player is to make morally grey decisions in the game, the player must first be informed of this and give

consent to engage with that sort of material. This means, there must be an understanding that the player is not the character, but is only exploring the worldview of that character. Different sides of a dilemma should then be represented by factions. These may be two companions that give the player their perspectives on a matter. Third, the motives of characters should be linked to a goal for the player. These can be points, or the promise of development of the story. Fourth, all characters should be humanized by having a personal history. Finally, the player should be encouraged to reflect on their actions both before, during, and after choices.

Visual design also has a huge impact on how humans read a person's morality (Bateman, Birk & Pradantyo 2021). As an example, Bateman et.al looks to real life where it can be seen how a person who is deemed attractive, either by normative standards or personal bias, is often considered more trustworthy compared to someone who is deemed ugly. Cults often use this to get members, where an attractive recruiter speaks to a potential prospect. These tactics are also used in sales. Visual design of a character can be used congruently, where design aspects that already have negative connotations are used to show that a character is evil. On the other hand, it can also be used incongruently, as a way to subvert non-threatening features to hide evil. Visual design can and should be used together with the steps suggested by Pluralsight (2015). The factions representing different sides can have facets that align with what they represent. A typical example is the devil and the angel on a character's shoulders when they are deliberating an evil/good dilemma. What Bateman et.al find in their study, however, is that many features used to represent evil traits are based on problematic stereotypes, and that using these substantiates these thoughts for users of the material. Racist stereotypes are very common, like dark skin (colourism) or large noses (antisemitism) indicating amorality. Due to apprehension to "ugliness", physical disability is also often demonized by society and media. Hodge, Taylor, and McAlenery (2019) did not find that real-life morality had a big impact on in-game action. This may clash with Bateman's text, as the basis of using congruent design facets is to bring the player's previous morality into the context to make them act accordingly when they for example meet a character that is intended to be read as amoral. Bond (2021) also finds that media representation of outgroups benefits pro-social behaviour in people outside that group. The reverse relationship is also true, something that propaganda material has always capitalized on. Furthermore, it can already be seen in Formosa's analysis of *Paper, Please* (2015) that games can dehumanize people through both narrative and mechanics. The game developer must be wary of how they construct evil to not authenticate harmful stereotypes by the usage of such tropes and character archetypes.

2.2 Engaging with player ethos

Consalvo, Busch, and Jong (2016) thought that "the magic circle" (the concept of leaving real-world notions at the door to properly engage with play) would allow players to explore actions they normally would not take. This was not entirely true, as almost everyone in their study wanted to play the role of hero. Interestingly, the participants were ready to commit atrocious acts if they aligned with what they thought was logical for the hero to do. To understand this, one can look at *How do video games elicit guilt in players? Linking character morality to guilt through a mediation analysis* (Ahn, Grizzard & Lee, 2021). Guilt is a high-level cognitive appraisal process where one attributes responsibility to oneself. Guilt can be lessened with external attributions, like arguing one was forced to commit an action. Game rules are often an external attribution. If the game requires killing civilians in the game world, the intensity of guilt will be lesser as that was the only way to move the game forward. To further explain players being likely to play as the hero, a contributing factor that was found to lessen guilt was that playing as someone who is perceived to be a high moral character (compared to playing as a low moral character) will do so. Ahn asks the question of whether a player would feel less guilty controlling an immoral character if the NPCs (non-player characters) were less moral. This is an aspect of this study's artefact. The notion may also be substantiated by Hodge, Taylor, and McAlenery (2019).

A trade-off for having an engaging game is that empathy may take a backseat, some argue. This is because engagement with a game means that the player is more likely to justify immoral actions. As Consalvo's study shows, this is even more true for the hero route of a game. In a utilitarian trolley-like dilemma, playing as the hero makes it even easier to sacrifice the one to save the many. The dissociation inherent in playing a game might also make more personal relationships in the dilemmas, like the one to be sacrificed in a trolley-like dilemma being the main character's sister, easier to make if the game engagement surpasses the player's empathy. This is of course with the caveat that it makes sense for the player character to make that choice.

Morality is a sort of expertise like many other aspects of human life. It is a set of social, cognitive, emotional, and perceptual skills. The four moral component model aims to map these skills out (Formosa, 2016).

(1) Moral Focus is about how committed one is to moral choices. With Moral Focus, humans can establish a strong moral identity that does not need external input for upkeep. It may include seeing oneself from the perspective of others, but consists explicitly of internal processes. In games, a moral alignment score can benefit this as it establishes the player's moral identity in the game world. Alternatively, it can be detrimental to have external input if the moral category

does not align with the player's morality, but they wish to keep up their score, therefore making some choices they do not agree with.

(2) Moral Sensitivity is about understanding morality in the real, external world. The concept includes both understanding the motivations of other people, but also one's own place among them. In games, this is important. In games that have a narrative focus, the choices can often telegraph their moral significance. In games that are built more around their mechanics, like *Paper, Please*, the player has to assess the moral situation themselves.

(3) Moral Judgement is what is most typically thought of as morality. Here is the understanding of moral philosophies and the tools used to reason about moral issues. The moral philosophies can be anything, like religious rules, military protocols, or just "common sense". Formosa (2016) argues that narrative games can have a wider breadth of moral dilemmas for the player to apply their Moral Judgment on than games mainly telling the story through mechanics. Mechanics limit the scope to the exact things the player can do, while for narrative-based games it's enough to simply *state* that there is a dilemma and for the player to pick how they would like to solve it.

(4) Finally, there is Moral Action. This component is about doing the moral thing despite it being difficult, or where committing immoral actions give better rewards. Therefore, it is about sticking to one's ideals and acting on them in a concrete way. Again, games structured with the basis of narrative before mechanics has some advantage in this area. It will be easier for the player to get what they want, as they only need to employ their moral skill, and not technical skills like eye-hand coordination. If the Moral Action must be done with mechanics, it requires both skill with the mechanics and with the moral components.

With this morality component model, some things can be confirmed. A morality meter system can aid the player's Moral Focus and establish a moral identity (Formosa, et.al, 2022) within the game world. It also allows keeping track of what faction alignment the player has if the narrative has different ones, which is often employed in games with moral choices (Pluralsight, 2015). However, Formosa (2022) highlights that it can take out much of the complexity in moral dilemmas. Narrative and mechanical approaches have different strengths and weaknesses. Focusing on the narrative makes the design of moral dilemmas easier for the developer. Flagging for what the choice's significance and potential consequences may be is more straightforward. This makes it easier for the player to perform the actions required of them to reach their goals. For mechanical approaches, the focus of the dilemmas must likely be smaller, and the tools to perform moral actions are fewer. It can however be used to continuously engage the player with moral decisions and constantly question their choices, also a component of composing games with a morality system (Pluralsight, 2015).

2.3 Psychology

A prominent theory about how humans make moral judgement is the dual-process theory (Greene, Nystrom, Engell, Darley, & Cohen, 2004). The two processes have different response times in the person making a judgement. The faster is called the affective type and favours deontological rules the person making a moral judgement has ingrained in them. It is unconscious and effortless. Areas of the brain that process emotions were the most active in both Greene et.al (2004) and a respondent reanalysis (McGuire, Langdon, Coltheart & Mackenzie, 2009). The other process is the cognitive one. This one is slow and takes effort. Mainly the dorsolateral prefrontal cortex is engaged during cognitive judgement. Humans have the most developed prefrontal cortex among all animals which is why humans can use these cognitive judgements to find the most utilitarian solution. This may be notable, as it shows that utilitarian ethics is only possible with access to a highly developed organism.

Greene's response (2009) to the respondent study (McGuire et.al, 2009) mostly defends his own study, but it does highlight a separate but related theory in moral philosophy. The personal/impersonal is another dualistic system. A personal moral dilemma is an action that could reasonably lead to the harm of another person. An impersonal moral dilemma is any conflict that does not bring harm to another person. Greene's study (2009) does confirm previous findings (Greene, 2003) that subjects with lower utilitarian thinking had quicker response time, ergo, deciding whether to reject a moral judgement in a short time, than the higher utilitarian subjects did about personal dilemmas. What these studies show is that most people employ emotional processes when making moral judgements about actions that cause harm to others. The cognitive processes, like utilitarian decisions, take a long time and is not as instinctual as emotional ones.

2.4 Theology

Jun's study (2005) compares the moral integrity of protestant Christians to atheists in South Korea. According to this study, the protestants did have 'higher morals'. However, an issue that can be seen is that it was based on Christian values. As an example, marriage fidelity. Although cheating in marriage is generally not accepted, it tends to be seen as worse by people of certain religious affiliations than how atheists see it. For atheists, it is a personal betrayal, while in Christianity, it is a spiritual one on top of the personal. The difference in moral integrity was

also much smaller than they initially thought it would be, which may suggest that moral judgement was not too impacted by religion.

Comparing how different religions see the concept of sin may be fruitful. Cohen and Rozin (2001) compare the mental states in judgement between Christianity and Judaism. They found a major difference. Jewish people do not put a lot of weight behind *thoughts* about immoral actions, instead opting to only judge oneself and others on real actions. Christians do however take thoughts about immoral actions into consideration. A thought is put on the same level of severity as the action itself. Notably, Jun's study (2005) asked people's opinions on immorality with no distinction between the desire of for example cheating and actually doing it. As they are both bad, according to Christians, the reason for omitting a distinction may be because there was none for the creator of the study. With Cohen's findings in mind, the dual-process theory (Greene, et.al, 2004) corroborates the Christian way of seeing sin to be that they see every moral dilemma as using cognitive processes.

Judaism instead means that it is impossible to reject a moral "action" such as having a bad thought through a cognitive process as it hasn't even been thought about yet. An explanation of the Christian way of seeing thoughts the same as action can be found in the meaning that they ascribe to life. Humans are put on Earth to essentially have their souls assessed and tested before the afterlife. Corporal life is only a stop on the way to eternal life, where good deeds accrue a way to heaven, and bad deeds may lead to hell. Christian actions/thoughts are always measured on if it gives reward or punishment. Seeing life as an exam in morality, it makes sense that everything, even thought, is seen as something that is judged. In Judaism however, life on Earth is more important. A rabbi (Freeman, 2020) explains it like this: "We're not in the business of getting to heaven. We're in the business of bringing heaven down to earth." Freeman goes deeper into this and explains that the reason that little is mentioned of heaven in the Tanach (Jewish scripture) is because the goal is to create paradise on Earth through millennia of collective labour by us humans. A tangible expectation of reward (or punishment) for what one does on Earth is also not important or present. Good is done because it is right. Here, too, some context can be seen for why Jewish people care more about actions than thoughts. Every action one takes is supposed to spread more good in the world. The concept of a Heaven above and Hell below is hardly even relevant in Judaism, while it is central to Christianity.

These texts may be used to inform the construction of moral judgement decisions used in the study. It is important to note that even similar religions, or people of the *same* religion, will have different interpretations of right and wrong. Jun's study (2005) has some issues, namely that there may have been a bias to prove that atheists are immoral compared to Christians. The fault within the questions, namely that the moral issues they measure are based on Christian

ideology, does show an important issue within how ethics function when compared to other systems. One will likely judge an issue from the morality one already has, and condemn others. If the study had a reverse bias, i.e. proving that atheists are more moral than Christians, the Christians would have been judged according to more secular values. One example would be the corporal punishment of children. This is almost universally condemned in secular ethics. However, a conservative Christian may look to certain passages to show that it is not only accepted, but encouraged.

"For the moment all discipline seems painful rather than pleasant, but later it yields the peaceful fruit of righteousness to those who have been trained by it."

(English Standard Version Bible, 2016, Hebrews. 12:17)

"Whoever spares the rod hates his son, but he who loves him is diligent to discipline him."

(English Standard Version Bible, 2016, Proverbs. 29:15)

"The rod and reproof give wisdom, but a child left to himself brings shame to his mother."

(English Standard Version Bible, 2016, Hebrews. 13:24)

A person outside this framework for ethics would likely proclaim this a call to parents to abuse their children. To the conservative Christian, it is not abuse at all if done with good intentions and without causing permanent physical harm. This good intention may be to instil discipline or respect towards them as a parent. A disconnection can be seen here. How can a religion that is non-violent encourage, from today's standards, child abuse even if that's not what they call it? However, the same kind of disconnect can be seen in non-religious settings. If a person robs another person, this is universally seen as a crime. Yet if a factory worker produces a product that is sold for 100 dollars, the company/boss collects 99 dollars, and the worker gets 1 dollar. The capitalist says; this is what the labour involved was valued at and is stipulated as such in the contract the worker signed. It is fair. The communist says; this is wage theft. The employer did nothing to produce that item. It would not exist without the worker. This is wage theft. The contract exploited the worker as their only option was accepting that 1 dollar, or none at all. The employer is worse than the thief. The disconnect here can be seen mainly in the capitalist. They see the personal transgression as a crime, but do not see the larger transgression (according to the communist) as a crime at all. Similarly, physical punishment of children for the conservative Christian is not a violent act in their eyes.

2.5 Benefits of morality systems in games

Ethics and morality have historically been largely informed by religious doctrines, and even the frameworks used in secular societies still have remnants of these deontological norms. Jun's study (2005) showed that the differences between moral integrity and even moral values in the atheists and Christian participants were quite similar. Whether or not humans can be considered to have an inherent baseline of morality, something that according to contemporary science is believed to be true (Greene, Nystrom, Engell, Darley, & Cohen, 2004) (McGuire, Langdon, Coltheart & Mackenzie, 2009), external sources will influence the further development of it. Formosa (2016) defined morality as a set of skills with four components. A skill is something that must be learnt or taught, and requires maintenance to be sustained and evolve. Religion has been central to forming people's ethical systems. In a secular society, however, this is done through education, and notably, through media. Bond (2019) saw how using parasocial relationships in media helped the participants of the study to better understand and establish support for outgroups like queer people.

When it comes to video games, the potential for honing and encouraging pro-social behaviour is even more prominent (Hodge, Taylor, & McAlenery 2019). As they are interactive, the player of a game should consider issues presented to them which allows for self-reflection of their values. If an immoral action is required or chosen, the player is confronted with that subject of morality (whether that is about kicking a puppy, or committing genocide against an aggressive alien species) regardless of what they end up doing in the game. Due to the potential of influencing morality, the way this is tackled in games must be considered with caution. If morality meters take the complexity out of a game (Formosa, 2022), the potential could not only be squandered, but also force a dualistic dichotomy onto the player. The rhetoric behind the choice would say that an action is either good, or it is evil. Locking into one alignment due to mechanical reasons (like locking rewards to "filling up" one meter) also has an underlying rhetoric that a person cannot change their opinion, or become a "better" or "worse" person. To build on this, some sources (Consalvo, Busch, & Jong 2016) (Ahn, Grizzard & Lee, 2021) showed that playing as a hero-type character made the players justify atrocious actions. A lack of complexity substantiates this, telling the player that a heroic character or organisation cannot do wrong simply because they are "the good guys". This can be particularly problematic in games dealing with topics such as war, like showing the US army's killing of civilians in a game in a sympathetic light, justifying their real-life atrocities.

3 Problem

Morality meters were mostly popularized by Bioware, with other developers following suit after the success of *Star Wars: Knights of the Old Republic* (2003) and *Mass Effect* (2007). Notable examples include *Infamous* (2009) and *Fallout 3* (2008). The system has been criticized for taking the complexity out of moral dilemmas and being a tool for the player to maximize their ability to reach desired outcomes. For some players, engagement may decrease if they are encouraged to pick the option that gives them the most points in the chosen alignment regardless of the narrative context or personal convictions. The morality meter system tends to be the most effective when there are two opposite sides, usually good/evil, or some variation of it. The player will likely believe both sides can be right, or equally wrong, depending on the dilemma. Another morality system is the more covert tracking of points without showing them to the player. This means it is a background process with no representation of it to influence the player's choices. The Japanese RPG series *Shin Megami Tensei* (1992) has been using this since the early 90s. Here, the player chooses between actions that align with *chaos* or *law*. Different endings can be achieved depending on what alignment the player has the most points in. The most ideal ending is the *neutral* one, which requires the player to balance their alignment. This has been notably difficult to achieve without using a guide or walkthrough as there is no consistent way to keep track of where they fall on the morality spectrum. However, if the player does not care about getting the neutral ending, this system is a good way to appraise one's morality in relation to actual dilemmas. The player is constantly asked to make difficult choices with no distinction between good and evil. Although the covert system in *Shin Megami Tensei* uses points to track alignment, these are not seen in the game. The reactions to the player's actions feel more dynamic as they can't simply look at their meter to know how they will be treated. Many games with similar systems also have reactions to specific events, regardless of the alignment. Both morality meters and covert morality systems have their strengths and weaknesses.

This study based itself on two main questions. The first is what system the player believes to be the most engaging to them while playing the game, something that affects their enjoyment of it. As video games are an entertainment medium, what will likely govern what systems to be used is the player's perceived enjoyment of it. If developers were to use a system that is generally loathed by their audience, this may have negative repercussions on sales. It is therefore pertinent to investigate this from a consumer perspective. The second question is what system is most useful for engaging the player with the material's complex moral questions. Will the player be more prone to keep to a certain moral alignment if the choices are flagged to give points in a moral category? Both sorts of systems, morality meters and covert feedback to moral choices, have their uses, it can be seen that one system will not work for any

type of game. This question also includes investigating if either system allows the player to be more flexible in their choices rather than sticking to what they started with if such is relevant to the player.

3.1 Method

This study is comparative. Two game artefacts were used, explicitly created for the study. Participants play one of these. At completion, they can fill out a survey. Below are more in-depth explanations of the aspects of the method.

3.1.1 Artefact

To answer the questions outlined, I wrote two versions of the same interactive text story (links to these can be found in Appendix B). An interactive text story is of indeterminate length and is identified by allowing the reader to influence the narrative by choosing between options given to them at certain points. I decided to use the tool Ink (Inkle) to write them, as it is a powerful game writing tool that allows for a lot of control. I based the construction of the morality systems employed therein on the overt morality meter from *Mass Effect* (a visual representation) and the covert morality alignment system used in *Shin Megami Tensei* (hidden from the player). Certain choices award the player a point in the choice's corresponding moral alignment. Due to the game being done entirely in text, it is very heavy on narrative which has both upsides and downsides compared to one that is heavy on mechanics (Formosa, 2016). Formosa says that mechanically led games can challenge the player's morality in a constant stream, while narrative games instead tend to do so during pivotal moments in the story. At first glance, it can seem like my artefact will challenge the player often, as choices in dialogue are constant. First and foremost, this is due to the brevity of the game. A short narrative game can afford to do so, whereas if such was done in a long narrative game, it would take up too many resources during production. Secondly, which is also something Formosa takes up, is that the breadth of choices is predetermined by me and presented to the player with no further options that could be more customized to the player's preferences. A mechanically led game would instead allow for more variation in the breadth of choices. This would not be desirable for the study. This is because I am measuring morality meters, which is a small number of alignments on ends with each other. And, as outlined in the problem section, the issues that for example Formosa (2016) finds with morality meters, like hampering complex morality, have been considered. Choosing one thing earlier will not lock the player out of other things down the line.

A large part of how society forms its perception of ethics is through religion. With this in mind, I chose to focus on two moral dilemmas. The rules of honouring one's parents whether or not the parent is abusive, as well as revenge, will be the main challenges the player is faced with. One type of choice will be about upholding those rules, while the other type is to go against them. These are not ascribed good or evil. The two dilemmas differ in severity and perceptions depending on the real-life moral system of the participant. The rule about honouring one's parents has a higher prevalence in people who recognise religious rules than someone who may be atheist (Jun, 2015). The dilemma of revenge is relevant for anyone, regardless of affiliation, or lack thereof, with a theological rule system. Both dilemmas will likely be judged cognitively rather than through an instinctual deontological rule. (Greene, et.al, 2004). Revenge, when felt to be justified, is *more* likely to be judged to be moral by most humans when only taking into consideration the affective and emotional judgement as revenge is a response to *another* personal moral dilemma. As another example, defending oneself against a robber is a dilemma directly related to the dilemma of robbery. This usually makes violence against the robber justified (McGuire, et.al, 2009). However, the context for knowing, cognitively, that revenge may be wrong can come from both a religious rule and/or a rule that has come from a secular community. The latter being especially true in a country like Sweden, where this study is based from.

On the other hand, respecting one's parents is more likely to be based on a religious rule, whereas in contemporary secular society the general development of cognitive rule is that one should only do so if the parents deserve it. Thereby, this will control if the participants in the explicitly labelled choices will change their alignment when they are presented with a different dilemma or if they will follow what they chose from the start to maximize points. The explanation for this control is that the typical atheist or religious person who does not strictly follow every rule may choose that not honouring an abusive parent is acceptable (in where they get points in one alignment), but may due to societal norms not accept revenge (which gives points in the other alignment). This control is necessary as having different choices that are likely to make the player keep to the same sort of alignment would not challenge the players who can see their score change to deviate from their initial alignment.

To further allow the player to make choices that they may truly believe in, their in-game avatar, ergo, the character from which they see the story, was not a hero figure. It was observed (Ahn, Grizzard & Lee, 2021) that playing a character perceived to have high morals lessened guilt in the player. The narrative context of a game will also affect the type of choices the player will be ready to choose. With the dilemmas being based on Abrahamic ethics (mostly Christian, as this is what I am the most familiar with), the game is set in hell. The player avatar was The Devil. This is meant to give the participant the means to judge without the moral scruples of real life and will detach the implications of choices from real-life judicial systems. Instead, the eternal

judgement of a soul is supposed to make the player consider what they *truly* think about these issues as no other consequences for the sinner's actions are possible afterwards. They will also not be influenced to act as a hero, as The Devil is not expected to act in a manner that is in line with higher morality. As ascertained by Consalvo, Busch, and Jong (2016), the player did not engage with the game freely if having the option to play as a hero figure. Based on both Consalvo (2016) and Ahn (2021), this should force the player to engage with the game and really consider their actions, as the "wrong" choice (to the player, if such exists) will elicit guilt. Sicart (2013) makes note of the concept of "wicked problems". These are problems that are similar to the typical trolley-type problem but are not meant to be consequentialist. Every solution is more or less "it's good enough", and never ideal. In games, they are meant to make the player feel complicit in the situation.

According to Pluralsight (2015), a key to writing games with moral dilemmas is to have factions. This allows for having voices to convey different perspectives on the issue, so that the player may be guided to reach their own conclusion. I therefore constructed two side characters to represent the two different alignments or ways of seeing situations. Bateman (2021) writes about how the visual design can say a lot about a character and their motivations. This has partly been subverted, as the character with traditionally evil traits (demonic) will advocate for the sinner the player will be judging, while the character with more traditionally good traits (angelic) will argue against their innocence. Problematic characterization of evil, meaning using human features that are stigmatized or based on discriminatory practices, has been avoided. The design of characters is further discussed in 4.2 Progression, in this paper.

The participant was presented with their final morality score at the end of the material regardless of version (overt or covert morality points), and was asked to input this into the survey. The only difference between the games is that in one the player can see how their score changes while playing, and in the other version they do not. The same choices were given the same score in both versions.

In summary, the artefact has been designed to allow encourage the participants to really consider their choices. Of note is not using a hero character to justify their morality, presenting every side as equal options, and subverting the design of the side characters from their typical personifications. One version of the artefact shows how moral alignment shifts according to player choice. The other does not show this.

3.1.2 Survey

To record information on how the participants play the artefact, they were asked to complete a survey form, the tool used being Google Forms. This tool was chosen as it is commonly used

to gather scientific or commercial data from respondents (Creswell, 2009, p. 149). Only their age group and potential religious affiliation were asked to identify them in terms of personal information, as these can have a large influence on their morality. Norms were different before, and differ from religion to religion. The subject matter of the artefact also ties in with religious themes, meaning this information may correlate with how a religious person operates in a theological context as opposed to a secular one. As outlined in *Survey Basics* (Phillips, et al., 2013), hypotheses like these inform the questions in surveys. Additionally, any stratification does not influence what version the person plays, as the study is distributed as either version A or B on different sites (Creswell, 2009, p. 155). Therefore, the participants were partly randomised in terms of what version they got. They did not get the option to choose the version either, unless they were to find two different posts. However, sharing on my personal Instagram for example, of course, meant that the participants were more likely to be the type of person that I am surrounded by, which would influence the identity background data. This will be further discussed in the discussion part of this paper.

Also relevant for background information is what types of games the respondent plays, if any. Sicart (2013) says that players, when interacting with a game, need accurate information about their state in the game. This can be highly influenced by the player's previous interaction with a genre. Sicart references Juul, who calls this "player repertoire". What this means is that it is easier to navigate a game if one has played something similar before, like understanding how choices may influence the game. Another example would be in horror games. A new player will always fall for jumpscares, while someone who knows the conventions of horror may guess what is to come before it happens. My artefact is made to be accessible. Anyone, even someone who hasn't played much besides *Mario Kart* with their children, should easily be able to understand how to navigate the game. The question about genre experiences is meant to check for this, especially while compared to other answers. The hypothesis (Phillips, et al., 2013) behind this is that someone who puts down genres thought of as "casual gaming" (gaming is not a prime interest to the respondent, and is simply a fun thing to engage with at times), will not be familiar with the conventions of choice based narrative stories. Of course, this is a subjective issue, as a casual gamer can play most types of games casually.

The participants were then asked about their thought processes and choices, as well as some questions regarding how they engage with moral choices in games if they do play games that deal with this. They were also asked for what reason they make choices in games in the sense whether they do it for rewards, and if they would ruin narrative integrity by looking up the most ideal course of action. This is paramount to the topic at hand, as a player who likes to play in the ideal way might prefer to know their "moral alignment" in the game. In essence, this is about if the player is reactive or reflective (Formosa, 2022) (Formosa, Ryan & Staines, 2016).

Due to the thought put into making the player feel complicit in the situation, rather than have them make calculated choices for the "best" solution, an important question asked in the survey was what made them choose as they did. Some answers include: "as I would myself", "as the main character", and "to make other characters like me". Finally, they are asked how they experienced the overt/covert system depending on the version they played, as well as the choice to say that they think they would have preferred the other. For a more detailed view of the questions and options in full, see Appendix A.

3.1.3 Distribution

The study was aimed mainly at people who are interested in narrative-focused games. Due to the religious aspects of the artefact, it would be desirable to have some of the participants who are/have been part of an Abrahamic religion to examine whether this could affect how they engage with moral dilemmas in games. This was however not the main goal. The study was distributed on pages concerning studies. It would have been shared on game-relevant threads on Reddit, but pages like the ones for such as *Mass Effect* or *Shin Megami Tensei* have rules that state that only material directly related to the media, like fanart or memes, may be posted. Seeing how players who specifically play these games that are the groundwork for this study would be of interest. A solution is that a question in the survey asks what type of games the respondent plays. Players who engage with a lot of RPG and narrative adventure games are especially interesting for this reason. A main source of distribution is via pages on Facebook and Reddit that are specifically for studies. This has the upside of gaining a wide array of people, and non-gamers. The downside is that the people who frequent these pages tend to be highly educated and academically adept. The study was also distributed via personal channels such as group chats or social media. Of course, people who have been in any way involved in helping with the construction of the study, such as proofreaders, were not included in the results or analysis.

3.1.4 Additional Issues

The main issue for this study was that completing it required a prolonged effort from the participants. The artefact takes between 15-25 minutes to read depending on the participant's reading speed. This was brought up to be an issue that potentially should be rectified if the response from the pilot study is negative. The survey should however only take 3-5 minutes unless the participant elects to write in custom responses, which may add some time. Additionally, the study is centred almost entirely on the participants' opinions, where these

responses are compared. It does not measure how someone's opinion on the systems being investigated may change when they are instead exposed to the other system. This could in fact be a strength, as it could show that *both* systems result in players wishing they were playing the other version. Due to being about hell and sinners, there could be topics that are not appropriate for this study. Dilemmas such as euthanasia, deadly experiments on humans for the greater good, and religiously motivated hate crimes were considered but rejected for it. The current dilemma does include physical and emotional abuse of a child, but is not detailed to such a degree that it may be too disturbing for most readers. A clear content warning is included at the beginning of the artefact, and will be communicated on every posting to advertise the study to discourage individuals who are sensitive to that material from engaging with it.

4 Implementation

4.1 Pilot Study

Both of the versions of the artefact (one showing changes to morality in choices, the other did not) were tested so that issues with both could be ascertained. The participants did not know which version they were to play beforehand. These tests were done by people who had a general idea of what the study is examining, as the topic of the study had come up in conversations beforehand. They had however not been directly involved in the process. The testing for the artefacts was focused on the flow of the story, the relevance of the moral dilemma, general enjoyment to play, and retention of player attention. The sample size for the pilot study was small, one person per version of the artefact, so there could be a bigger focus on qualitative interviews. No substantial issues were found with the artefact. During informal interviews (Creswell, 2009, p 173-193) with the respondents, some things were highlighted which will be discussed henceforth.

Prior to the pilot study, I thought that most players would be very skewed towards siding with the character representing a more secular and victim-focused philosophy over the character that represents something more in line with conservative religion. Part of this is because the study is based in Sweden, a country that is generally quite atheist, or is open to viewing religious doctrine through a critical eye. The views expressed in for example Jun's study (2005), where there is a clear divide in the South Korean culture when it comes to how morality through the lens of a protestant vs atheist morality, is not as strongly prevalent here. Through personal experience, even people who are part of a Christian denomination or sect tend to hold secular beliefs about such things as LGBTQ and women's rights, especially so with younger generations. This does show a bias in me, as I ascribed a judgmental view on the issues presented within the artefact to religious people while taking for granted that a non-religious person would be less judgemental. However, one of the respondents, who is not and has not had any affiliation with a religious establishment, almost exclusively sided with the judgemental alignment, in other words, the alignment based on Christian morality. The explanation for this was that *a:* however much someone has hurt you, their life is worth more than retribution, and *b:* they acted like they thought The Devil, the main character of the story, would. This is in opposition to the other pilot tester, who does come from a religious background (Islam, however, they are no longer part of any religion), who exclusively sided with the non-judgemental alignment, or what is written to represent a secular morality. The differences in how the respondents played tell us that the issues presented in the story are nuanced enough to not necessarily be predictable on account of past religious experiences. The

sample size for the pilot study was small however, so the question of religion with the potential correlation to in-game choices was kept for the study.

Furthermore, there was a concern that the study would be considered too long to complete. The pilot study respondents did however consider it to be an acceptable length. The respondents took about fifteen minutes to complete playing the artefact while playing at a leisurely pace. As it includes playing a game, the length was expected and the story was engaging enough to keep their interest. Both respondents commented that although the story handles difficult topics with sides that are "extreme opposites", all perspectives made sense and were appealing to some extent. One respondent noted that it did not feel "preachy", nor that it was trying to manipulate the player into choosing according to the writer's bias. These things are important to know, as the point of the study is to see the player's level of engagement with the game's morality.

One respondent, for the question: "When you play games with choices that deal with dilemmas of right and wrong, or some variation of that (Mass Effect, Undertale, Bioshock), how do you engage with the game?" answered chose the option:

"I try to get the best rewards regardless of if some choices do not align with my ideals, and may look up how to do it."

Pilot study, respondent #2

This was also the player who had full points in the alignment represented by The Advocate, who speaks for the in-game sinner's innocence. They also responded that they saw how points changed with choices. It is unclear if the exclusivity with this alignment is because they agreed/disagreed with the choices, or if it was due to points. The closest hypothesis would be that it was due to agreeing with that alignment, as for the question regarding what motivated their choices, they did not choose the option "maximize points". With the main study, a better correlation was thought to be able to be gleaned whether the type of player (reactive, wanting to maximize points) always chooses according to one alignment exclusively, or if this was an outlier.

4.2 Progression

An issue found within the survey through the pilot test was that there was no branch diverting depending on what version the respondent had played. The structure of the survey was instead "if you saw how your score changes depending on choices, answer this question. If not, skip this question." This was confusing and cumbersome. The survey structure was instead changed according to this feedback. In the final rendition of the survey, there was only a question about

whether or not their score change was visible in choices or not, whereupon the respondent only saw relevant questions to their version of the artefact if this question was answered.

No substantial changes have been made throughout the work with the artefact. The lack of changes was justified by the feedback from the pilot testers, where no negative remarks were given about it. Most thought processes have been outlined in the method section (3.1.1 Artefact). I shall go further into some of these points.

Essential to the system of choosing differing alignments, there had to be two characters to represent these. To subvert expectations, the character with an "evil" design, resembling a demon, was the one to advocate for the sinner's innocence. The angelic, or "good", was instead the one to argue against them. Visual designs exist for these characters, but were not used in the artefact as they did not fit in with the interactive text story. Only The Devil is shown to the respondent as the game icon on itch.io. For the sake of this study, pictures are available below (Figure 1, Figure 2, Figure 3). Descriptions in the artefact are based on these designs.

Figure 1. Design of The Devil character from the artefact.

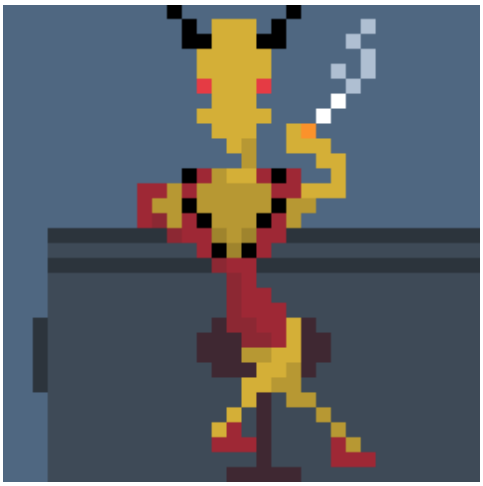


Figure 2. Design of The Advocate character from the artefact.

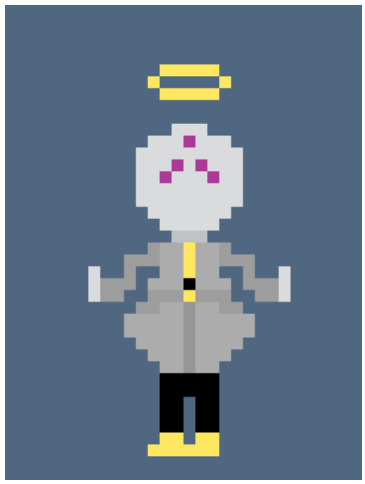


Figure 3. Design of The Ordinance character from the artefact.

As the choices cannot really be categorized as “good” or “evil”, I decided to simply name them after the character that represents the ideas presented within. This also encourages the player to stay engaged with the dialogue. This is not only true for the version where they cannot see score changes, but also for the one with scores, as the name of the category itself may not be sufficient to explain exactly what the choice would entail in terms of moral philosophy. Before any such choices were presented however, the characters were introduced with dialogue that exemplifies their morality so that the player may still have an idea of what, to a start, they want to align with. This also sets up the categories, where The Advocate shows their anti-establishment ideology and rebellious nature (what the alignment “Advocate” represents), and The Ordinance shows their prevalence for upholding the status quo and rule systems (what the alignment “Ordinance” represents). The visual aid in one version of the artefact is also presented in a simple manner, and is meant to be intrusive to challenge the player in how or if they will use that information. Figure 4 and Figure 5 show the difference in how it looks to the player during dialogue choices.

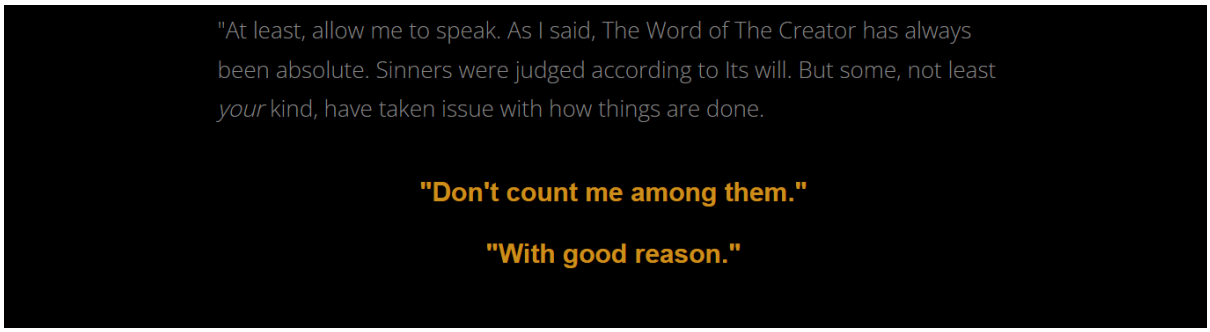


Figure 4. Dialogue choice without showing how alignment score changes.

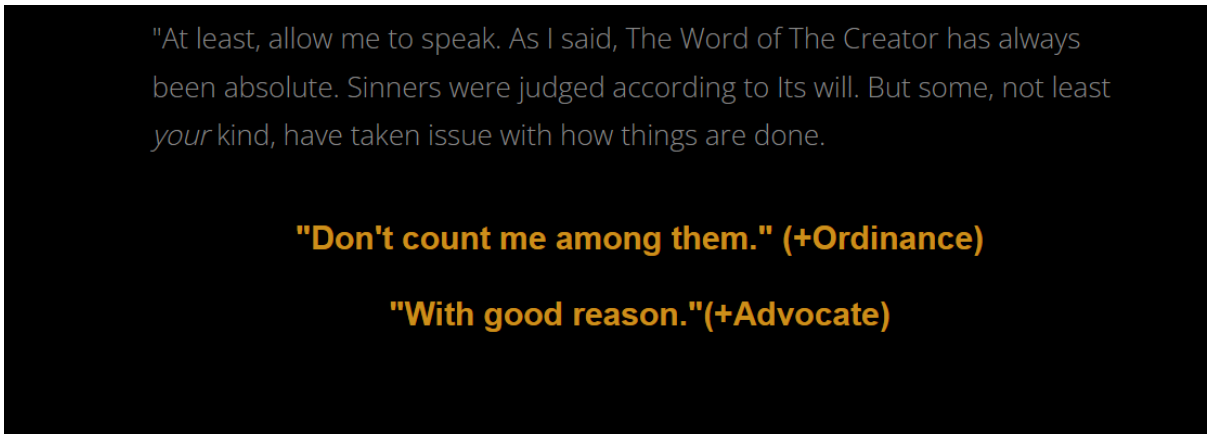


Figure 5. Dialogue choice showing how alignment score changes.

Many things have informed the design of the artefact, and the basis of the study. Notable is the discourse surrounding morality systems in interactive media (Formosa, Ryan & Staines, 2016) (Formosa, et.al. 2022) and the effects it can have both on gameplay, engagement, and the player's development of a moral identity, topics discussed in various of the material used for this paper (Formosa, 2016) (Consalvo, Busch & Jong, 2016) (Ahn, Grizzard & Lee, 2021).
The discourse is often centred around morality meters, the way of measuring morality in a way that is supposed to be objective. Of course, this is often only objective in a game design way, meaning that taking an action results in a point for the category the game designers decided represented that choice. In some games, different actions give a bigger number, where they have decided that, for example, this action was so evil it nets the player 30 evil points instead of 1. As stated by Formosa (2016)(2022), to accomplish the moral meters, morality is often simplified, which can be seen in games like *Mass Effect* or *Infamous*. In the case of *Infamous*, the case is essentially good or evil to the extent where they are caricatures of the two.

Mass Effect has the moral flavours "Paragon" and "Renegade". These are meant to represent whether the player always try to help others, or do whatever they must for the bigger picture. In some points in the story the player is confronted with a clear dilemma, like sacrificing billions for their goals. The options here are essentially; do you want to do it like an asshole or pretend to be nice about it? Overall in those games, no matter one's alignment, one will be responsible for innumerable deaths and have committed numerous war crimes. This is definitely interesting, but the game rarely actually questions the player on this. The "good" characters in the game always go to great lengths to justify the player's actions, while the characters that do oppose the player's genocidal rampage are deemed weak or "bad". This is true for *both* moral alignments. I do think that Formosa (2022) may be onto something when he wrote that morality meters tend to take away the nuance from morality, especially when it simply measures what flavour of a similar thing one is. It is also, as seen in *Mass Effect*, detrimental to a game that touts complex morality as its selling point where all options are

informed by the creators' bias. No matter what the player chooses, a pro-military agenda permeates every action. As mentioned earlier, the respondents in the pilot test for this study mentioned that the artefact did not feel preachy or biased to one alignment, regardless of whichever side they chose. This was something I was very careful with, as I did have a bias to one side. The way *Mass Effect*'s morality is constructed is what I had in mind when creating my own as I wanted to avoid the same pitfall of having a particular agenda permeating both sides.

On the topic of *Mass Effect* and morality meters, the games' biggest criticism is that the morality system forces the player to stick to one side throughout. Some big choices near the end of the games are locked behind filling out the entire meter in one alignment, and the ideal dialogue choices in the bulk of the game are often greyed out unless the player has stayed relatively consistent in their choices. In the artefact, I decided that even if the player chooses only one alignment throughout, they can still make their final decision regardless of their score. In other words, they can have picked only choices representing alignment B, but make the final choice that is in line with alignment A's morality. One explanation for this is that the player is encouraged to change their opinion on the dilemma throughout the game. Another is to preserve the complexity. One can think that yes, this character has committed wrongs, but they still do not deserve to be judged to hell for it. Differences in dialogue are closer tied to specific choices that were made earlier in the game (some choices are tracked with variables, and dialogue further on is based on this, not the score). In a larger game, it may be pertinent to tie some dialogue choices to unlock only at a certain alignment score, like in *Mass Effect*, for the purposes of looking into this aspect of morality meters. However, this was not relevant for the study, as the player only needs to think that their score might matter to pursue a certain alignment if they so choose, and it could broaden the scope of the study too much.

The other type of morality that Formosa (2016) discusses is the type found in *Paper, Please*. This morality does not use any type of direct scoring system, and the morality is not predetermined by the creator. This is a morality based on the game's mechanics, rather than its narrative. This type of system would be good to use for the moral dilemmas I have used, as it retains the complexity of morality by allowing the player to really weigh the benefits of their actions, while engaging their ethics. *Paper, Please* excels at asking the question; can you do what's wrong if it's to the player's benefit? They are almost forced to do what's wrong to keep the game going, as it is notoriously difficult to do what is right and without a game over screen. However, this system would be *too* complex for the artefact, which is after all a narrative game that consists mostly of text. I still brought with me the complexity that *Paper, Please* and games like it strive for. However, reading about it inspired me to look at games like *Shin Megami Tensei*, which uses a hidden morality system that still tracks the player, but one can go the whole game without even knowing that such is the case.

In summary, the pilot study mostly found issues within the survey concerning the structure. This was remedied to better guide the participant towards the correct questions, meaning, questions relevant to the version of the artefact they played. Concerning the artefact, the pilot study found that the length was acceptable. During the construction of the artefact, I based much of the moral alignment system on using the issues found within morality meters and morality told through mechanics, while also finding a moral system that tracks the player while not showing this to the player. The latter is believed to be a good middle ground for still using morality meters, while retaining the complexity of the moral dilemmas present within the text. The versions of the artefacts would therefore be one that functions similar to *Mass Effect*, where it is easy to track what one's alignment is, and what morality is ascribed to choices. The other version is like *Shin Megami Tensei*, still ascribing morality, but not showing this to the player. The alignments here are also not "good" or "evil", or a flavour of this (as in *Mass Effect's* "Paragon"/"Renegade"). Instead, I have constructed morality to be different views of what is right and wrong. The choice of using an angel to represent the "harsher" side and a demon to represent the more "empathetic" side may also further subvert notions of good and evil, urging the player to think critically about their choices.

5 Evaluation

5.1 The Study

The study received responses from 23 people. However, two of these were part of the pilot study, and therefore are not going to be included in the results as the survey changed slightly due to their feedback. Additionally, one other respondent must be removed due to giving non-serious answers. The explanation for this is that the individual did not read the artefact, and in the survey only said that they do not play anything but VR games, ostensibly to trivialize games that are not part of that medium. Therefore, the response is neither relevant nor beneficial for the study topic. Of the 20 people that will be used, 14 played the version of the artefact that does not show moral alignment changes with choices. 5 people responded that they played the version of the artefact where these changes were shown. The final respondent says that they do not know which version they played. This response will be kept, as their responses are still relevant in the part of the survey not specific to a version of the artefact. It is also not considered a non-serious response, as answers to player background suggest that the person is simply not used to navigating game spaces.

The respondents were relatively well dispositioned in terms of age, although the biggest part was within the 18-25 age range (figure 6). In terms of identity, the question of religion was also brought up (figure 7). The vast majority identified themselves as atheist and/or agnostic. Two people are Christians, one person is Muslim, and one is part of a religion that is non-Abrahamic (not specified which). One person also responded that they do not know, which will be counted as separate to agnostic, but also not religious. Finally, when it comes to video game genres, the vast majority said that their topmost played genres are RPGs (like *Mass Effect)*, story-driven games (like *Life is Strange)*, and simulation games (like *Minecraft*) (figure 8). Otherwise, additional genres tended to be tertiary to these ones, as in, they are the common denominators, and the rest is up to more individual preferences. The response suggests that most of the participants have some familiarity with story-centric games.

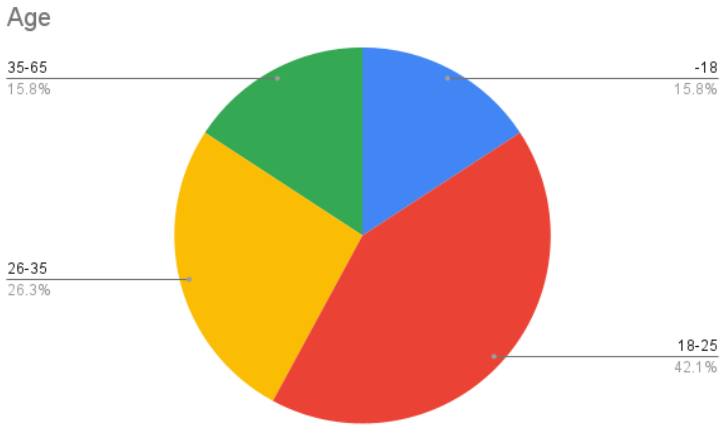


Figure 6: Age groups

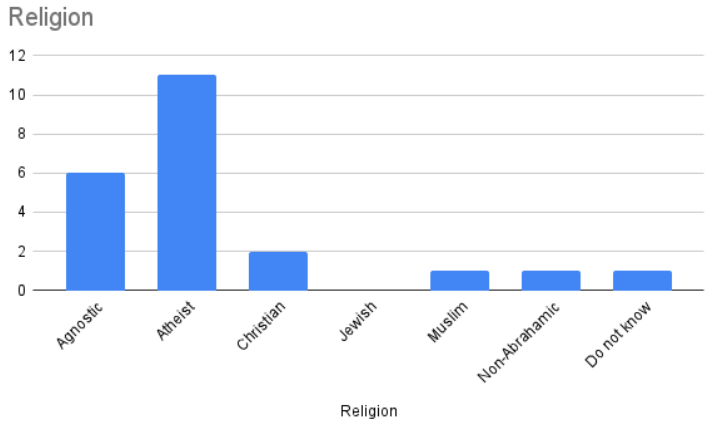


Figure 7: Religion

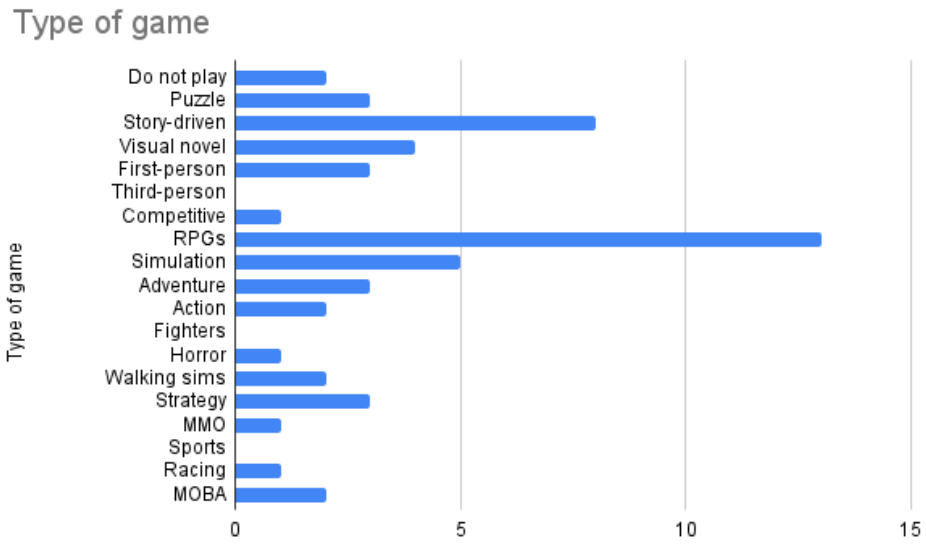


Figure 8: Type of games participants play

The final question before really going into the participants' experiences and thought processes with the artefact pertains to their moral alignment score. 18 people answered this question (it

was not mandatory to do so, in case they had closed the game window and therefore would not be able to continue the survey.) Almost all of them leaned heavily into "Advocate", which represents a more forgiving attitude towards people who have done wrong. It can also be seen as an anti-establishment alignment. The other alignment, Ordinance, is instead about upholding law and systems, no matter the person or the circumstances. Only two people could be considered to have relatively major sympathies towards this alignment. One had a score (Advocate:Ordinance) of 5:5, and the other 6:4. As outliers, these will be examined more in-depth further on.

Two questions ask about how the participants play games. The questions are in full: "What made you pick the choices you did?" (figure 9) and "When you play games with choices that deal with dilemmas of right and wrong, or some variation of that (Mass Effect, Undertale, Bioshock), how do you engage with the game?" (figure 10). The answers for the second question take into account that some may not play games, or games like them. However, no respondent took advantage of the "I do not play games" option, and it was answered by all 20 participants just like the first question. Nobody chose the option that they were trying to maximize their score. The majority instead said that they made decisions based on how they would act themselves, and what they thought were the most "good" (as in, good vs evil). From the other question, it can see that the majority are reflective players, which is substantiated by the self-insert play (self-insert play: make choices based on their own morality, or a projection of themselves onto the main character). Only 4 participants said that they tend to play for reward and would look up how to do so, while the rest of the people who chose given answers responded that they like to take things as they come.

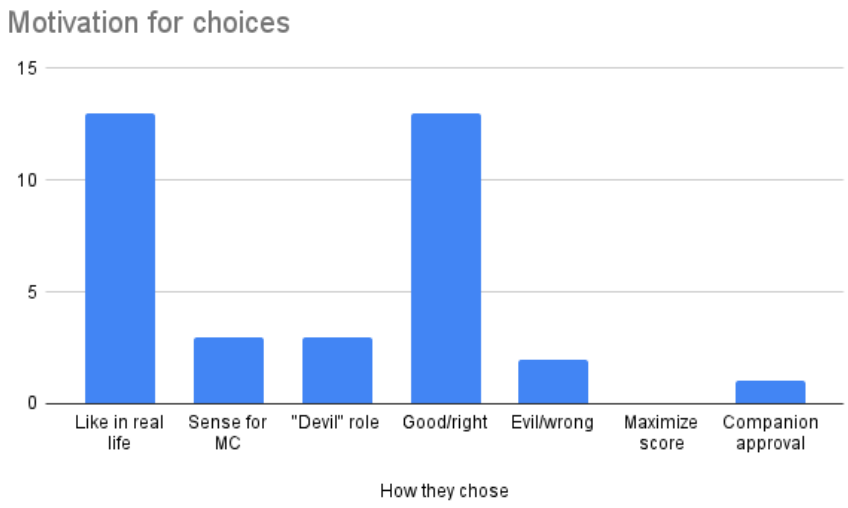


Figure 9: How participants made choices

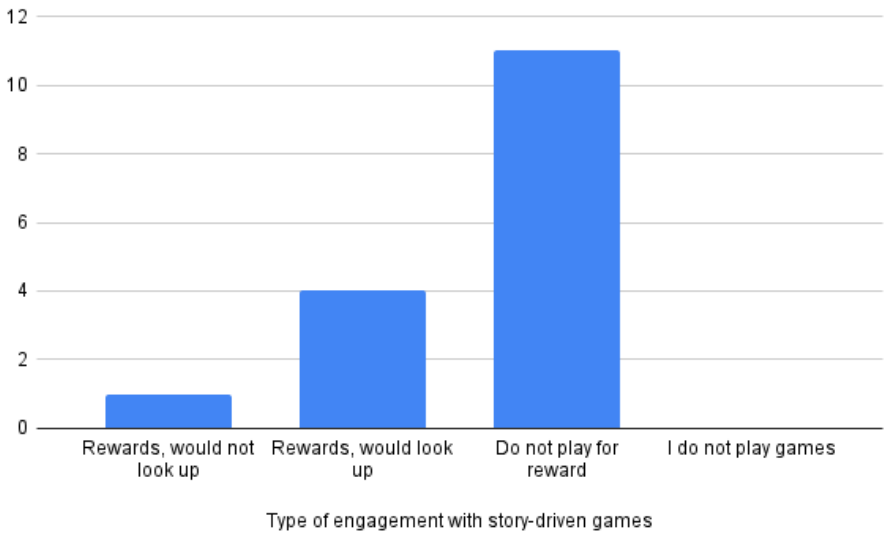


Figure 10: How participants engage with narrative games

The latter question also got write-in answers.

"On a first playthrough, I pick the choices that I would make in real-life (and not caring about the reward), but for other playthroughs I pick the other choices and take the reward as I go."

-Anonymous participant, response 3.

"On my first run of games, I will try to play according to my individual morality. On later runs, I play according to reward or other roleplays."

-Anonymous participant, response 7.

"Generally on my first playthrough, I will engage based on my real beliefs, but on following playthroughs will play to maximize rewards."

-Anonymous participant, response 9.

"I usually choose what I would in real life, sometimes I get in character if I'm not on my first playthrough."

-Anonymous participant, response 16.

These responses all have the same meaning in regards to how they play games. The two parts of these answers are: playing as themselves on their first playthrough and taking things as they come without looking to maximize reward or influence the story according to how they want it to go; on subsequent playthroughs, they may however act in a roleplaying manner (either as the character, or a character they have made themselves if the game allows for customization),

or to get the best rewards. This means that this type of player prefers reflective play, but also enjoys reactive play when they have gotten the "real" experience.

Further into the question of how the participants made choices, which allowed for multiple options, more can be seen about the intersections of how people play. Looking at them, it can be seen that the combination of "I chose how I would act in real life" and "I chose what I thought to be good/right" was relatively common, and definitely the most common out of all other combinations. Nine people had these two responses at the same time. From this, it can be inferred that it is typical that people consider themselves to be good people, and aim to navigate spaces by spreading that good to the best of their ability. One type of response may be confusing, however. Two of the people with the aforementioned combination had the additional response "I chose what I thought to be evil/wrong". Looking at the responses, it was unclear how this question was interpreted by these respondents. Did they seek out both good and evil choices, or did they choose options that they are self-aware about being wrong, even while choosing them? A choice can also be simultaneously right and wrong depending on perspective and context.

Additionally, two other options were also chosen by some people. These are "I acted as what made sense for the "The Devil" role", and "I chose what made sense for the main character". Despite these answers being very similar in meaning, none of the six participants who responded with either of these actually chose them together. They were instead an individual answer, or it was in combination with another option. This suggests that the main character, and the gestalt of The Devil, are not necessarily seen as being the same. This is interesting as the main character *is* The Devil. While nothing can be said conclusively about this, I do see two possibilities. The first is simply that if they chose one, they did not feel like they had to bother with the other. Another, in the case of choosing that they acted as the main character and not the role of Devil, may be that the gestalt of The Devil (the picture people have of the character, based on all the pieces of information from a myriad of sources people possess of it) does not fit the characterization of them in the artefact. With the opposite, choosing that they acted according to the role of The Devil and not the main character, could be that an antagonistic character, like The Devil, tends to be seen as in many cultures, cannot be the main character.

Furthermore, no participants chose that they played to maximize their score, and only one chose that they picked choices according to what characters would like them more. This was not entirely expected, as the characters were constructed to be engaging characters that are there to guide the player to make choices based partly on characterization. It may however suggest that the players were more concerned with engaging with the moral dilemmas rather than playing favourites with the characters, or the two alignment characters weren't likeable themselves.

All this suggests that the participants, regardless of past experiences with games, were engaging in reflective play. They also tended to not seek external validation from others, at least in the context of this study. The participants can therefore be thought to have strong moral focus, meaning that they know their morality and can navigate dilemmas without being affected by, or even need, the influence of others.

Looking for trends in the way the participants engaged with the game depending on the version they played, one can look to the responses for their moral judgement about the characters in the story. Four questions pertaining to this were asked. One is if the sinner should have treated her mother better (figure 11), whether she was motivated by revenge (figure 12), if she was responsible for the death of her mother (figure 13), and finally how the respondents' opinions changed throughout the story (figure 14).

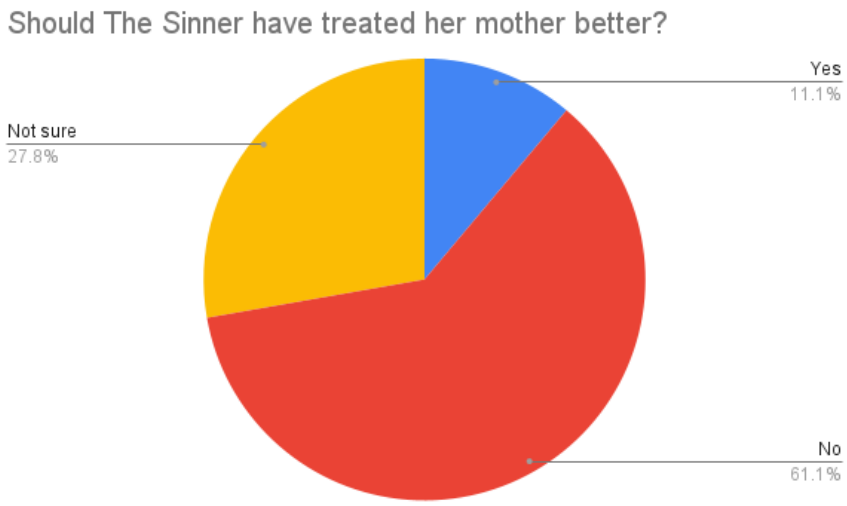


Figure 11: Opinion on if sinner should have treated their mother better

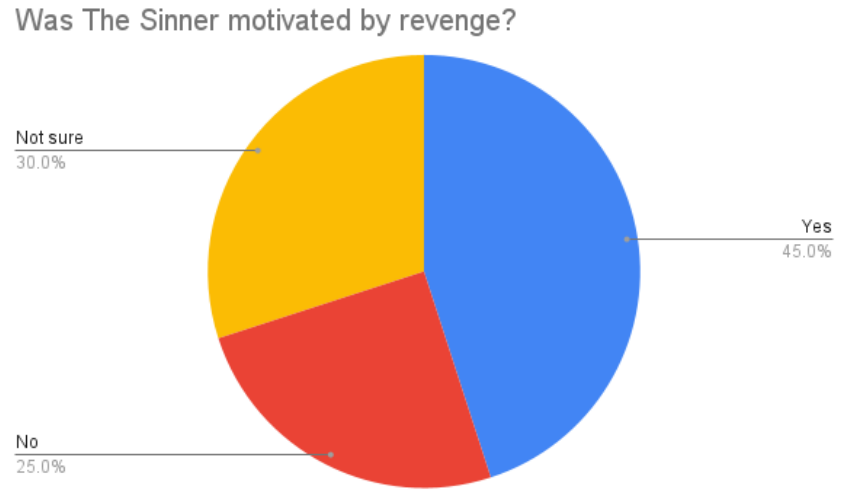


Figure 12: Opinion on revenge as motivation

Figure 13: Opinion on responsibility of death

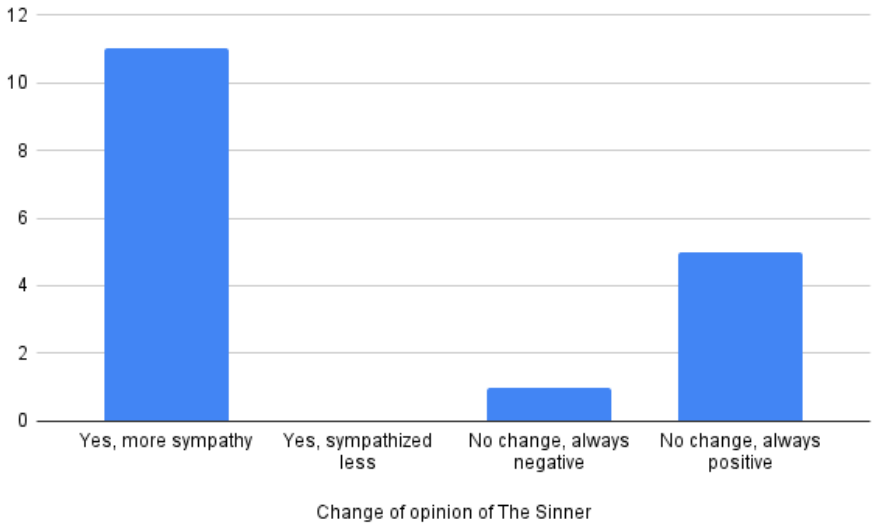


Figure 14: Change of opinion throughout play

For the question about whether she should have treated her mother better, the general consensus is no. Only two people gave a resolute yes. Write-in answers suggest that this was a difficult decision.

"I think, in that one instance at the end, she should have. I don't think it makes her a sinner. To judge a situation like that in solely black or white is - you miss the nuances of things. But personally, I think she should have at least done something. If you're not that kind of person, though, I'd understand too. I'm much too empathetic a person to people who might not deserve it. And I understand that anger. It's heartbreaking, and ugly, and sometimes you just have to let that go. Even if it means cutting yourself off from family completely, or helping them out briefly as a person."

-Anonymous participant, response 4

This person says that the sinner should have helped, but that a situation like this isn't black and white. Even if one should do something, it doesn't necessarily make them a bad person if one chooses inaction, to cut oneself off completely from the person who has hurt them. Another participant seems to agree:

"Yes, but it's understandable why she didn't and I don't think it makes her a bad person."

-Anonymous participant, response 10

Another write-in response to this question is harsher in their judgement.

"*I fully understand her frustration because a mother should never treat her child like that. But wishing death or harm will not make anything better. The sinner does not owe the mother anything. But she could have avoided saying what she did."*

-Anonymous participant, response 11

This response is interesting as it touches upon the issue of thought/action that is a point of contention in Judaism and Christianity (Cohen, Rozin, 2001). This respondent's response suggests that the sinner's hatred of her mother would be ok as long as she did not put it into words, ie, not tell her mother that she wishes harm on her. Cross-referencing with the personal responses, this person does identify as an atheist. When it comes to religion, this would be in line with Jewish morality, rather than Christian. As the artefact is based on this kind of moral dilemma, it is interesting to see someone putting it into words without having a direct background with a religion where this part of the ruleset.

The response to whether she was motivated by revenge was more split. Nine people said yes, five said no, and six were not sure what to believe. The write-in-answers highlight her simply not helping/being apathetic towards the situation, questioning whether or not this can be considered revenge.

"*I guess she was, but does refusing to help even classify as revenge?*"

-Anonymous participant, response 14

"*I believe what she had said: there was likely vengeance in her heart, but not in her actions. Her actions were defense (suitcase) and simply remaining estranged (bile)*"

-Anonymous participant, response 7

"*Perhaps only a little, greater so by apathy*."

Anonymous participant, response 9

"*The 'sinner' was simply indifferent - that's not revenge, maybe even arguably worse.*"

Anonymous participant, response 12

Finally, the question regarding whether she is responsible for the death of her mother seems to reflect the scores of moral alignment. Only one person put down for the given "yes" response, with some write-in answers also reflecting on it.

"Since the sinner could have prevented it, yes. However, the mother is also responsible for her own death."

-Anonymous participant, response 9

This person's answer, despite the brevity, could be quite nuanced. They acknowledge that the sinner could have prevented the death, but also says that it isn't without preexisting conditions. The mother, by being a decent parent, would have prevented her own death by having a daughter who loved her. Every happenstance is connected, and the circumstances around our actions are usually influenced by other people's actions, even if we can't see them. Another participant reflects further:

"*In a way, she is responsible. She left her when her mother reached out for help. But we can't tell the future. Her mother had hurt her irreparably and the daughter, understandably, wanted nothing to do with her. The anger that comes from a parent treating you that way is not something that just goes away. Neither does the regret or sadness of it all. You may decide to forgive, but never forget.*"

-Anonymous participant, response 4

Yet again there is an interconnection of other people's actions with one's own. Of course, one cannot know what one do will make other people do. The mother abused her daughter, which in turn resulted in her own death. The response validates this cause and effect, and is the participant's justification for why the sinner is not responsible. Instead, what happened was the mother's decision, and the daughter's one is just the consequence.

Interestingly, two people said something that is in the same line as each other.

"*not at all, if this was the case each person would be responsible for death of hundereds of others*"

-Anonymous participant, response 14

"*Yes, but no more than she was for all suffering in the world that she didn't try to prevent.*"

-Anonymous participant, response 16

The two responses can be said to divorce the notion of the importance of morality depending on how close someone is to us socially. These responses beg the question; is it worse to not help a family member rather than, say, a random stranger? A person is shunned for not helping their mother and the responsibility to do so is put on child. Yet if one was to walk past a homeless person in the winter, deciding not to bring them into one's home, nobody would ever blame the person for when this homeless person dies from freezing to death. From a purely moral standpoint, it doesn't matter what a person's relationship is to oneself, wrong is wrong, and right is right. Of course, people put our their social values onto morality, which is why one is acceptable and the other is not. In the situation of the artefact, the sinner has made it clear that her mother means just as much to her as that homeless person, meaning, she should be judged on those terms. But there is an intrinsic value put on the relationship to the people who share the same DNA as oneself and it is not something that most people wouldn't take into consideration.

To summarize the response to these three questions, the participants generally felt that the sinner may have been motivated by revenge, but that doing nothing to help her abusive mother is justified, as she did not actively attempt to cause her mother's death.

The final question, "Do you feel like your opinion of The Sinner changed throughout the story?", follows the trend of being generally favourable towards the character. Sixteen people were positive, where five had that impression from the start, and eleven changed their opinion according to new information being presented to them (this question allowed for multiple answers). Only one person had a negative impression from the start that did not change it, and nobody sympathized less as the story moved along. The three participants tried to stay more impartial, as their opinion were neutral. One person wrote in an answer:

"*I considered the situations in theabstract and did not feel either sympathyor antipathy for The Sinner.*"

-Anonymous participant, response 18

When looking at this person's personal responses, they are in strong favour of the sinner's innocence, as they said no to all the questions about if they were in the wrong, and their score was A:9 O:1. They also chose according to what they thought was good/right.

Having gone through the responses for the general part, one can now move on to the parts specific to the versions. To reiterate, the amount of people is not symmetrical, where version B, not seeing how the score changes, has more responses than version A, where one saw how the score changed.

The first question asks how the type of score measurement affected them. For version A, three people said that they would have preferred not to see the score change (figure 15). People still felt like it had positives, especially in the way that it helped them follow the alignment of their choice. Nobody who answered this question considered this type of score to pressure them into sticking to one side. Of course, the sample size for this version was rather small, so others may have felt differently.

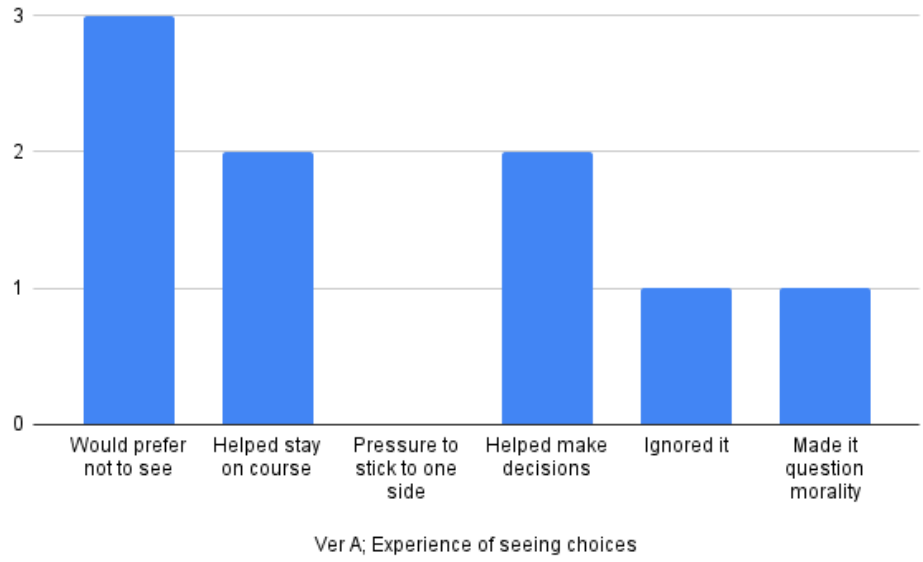


Figure 15: Group Version A, Experience with artefact

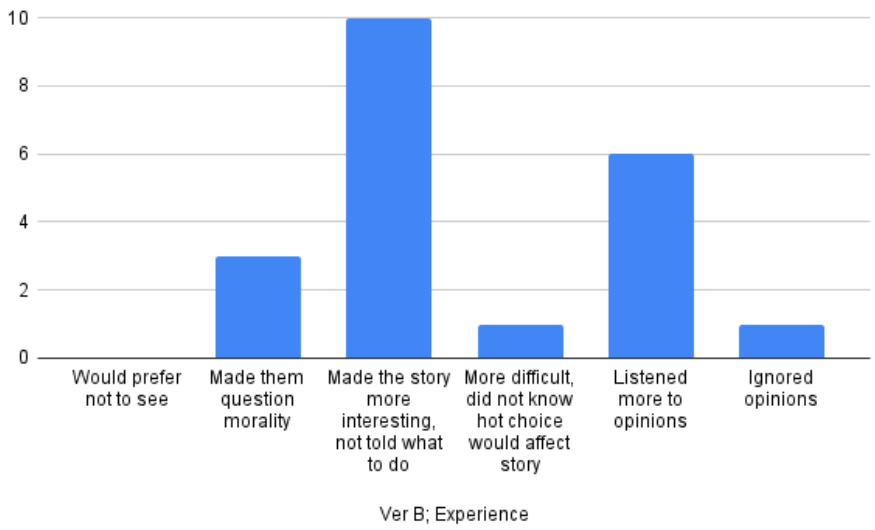


Figure 16: Group Version B, Experience with artefact

Version B had a similar question (figure 16). People generally considered this version to allow for freedom in how they engaged with the narrative, with many people answering that they enjoyed not being told what to do, and some questioned their morality due to this, as their standpoint was not categorized for them. This may also be why many listened to other characters' opinions. None answered that they would have preferred to see the score changes, while one says it made it difficult to make choices as they did not know how it might affect the story. However, two write-in answers note that they found it clear how the choices would affect the story and characters, meaning that the presentation of the story may be too transparent to accurately measure how not seeing this score change could confuse a player.

The other two questions for both versions ask about if the visibility/lack of visibility of score made them engage more with the story, and if they stuck to one alignment side. For engagement, the answers were opposites. For the people who could see the score changes, the majority said that it did not make them engage more with the story, whereas the majority of people who could not see their score change felt like it made them engage more. Additionally, the participants in version A were more prone to change their alignment as new information was presented to them. The rest in version A stuck to one side, and none had mixed alignment. For version B, the vast majority was on one side, with two people mixing their choices, and one who changed their alignment due to new information.

5.2 Analysis

Looking at the demographics of the study, it can be established that it is generally of a younger group, meaning up to the age of 35 years old. Additionally, many respondents circled in RPGs as commonly played genres of games they play, suggesting a high player repertoire, a concept discussed in 3.1.2 The Survey, taken from Sicart and Juul (2013). In brief, these players have knowledge of similar games and can easily pick up how the artefact worked.

The demographic of participants is also generally non-religious, with some representation from religious people. Many of the participants were sourced from Sweden or Finland, a generally atheist region. From looking at their answers, it can be theorized that secular morality is present in most of the participants, atheist or not. This may be due to several reasons. The first and foremost is that they are of a younger age. Conservatism tends to be more prevalent in older generations, and young people tend to be more receptive to contemporary ideas. A large part of the participants is also sourced from either spaces of higher education or my own social circle. Secular and progressive ways of thinking are generally more common in higher education, in both atheist-leaning and religious countries. Of the people in my social circle who participated, many are of the political left. These aspects likely contribute to players of the artefact leaning heavily into being amenable to the plight of someone who has been victimized by their parent. As one of the main things being investigated is if morality being flagged in-game or not will contribute to sticking to a certain alignment, this is relevant as an explanation to the high percentage that favoured the Advocate alignment.

However, one of the bigger outliers in terms of score ratio (5:5), answered that they are Christian and between the ages of 18-25. Such cases show that it is not set in stone that age necessarily prevents a more judgmental morality even if religious. Similarly, the person with a ratio of 6:4, still in favour of the Advocate, but higher Ordinance score than others, is in the same age group, and is not religious. This shows that being non-religious does not necessarily mean having non-judgmental morality.

Furthermore, the most important question for the aforementioned topic, flagging of choices contributing to keeping to an alignment, is found in the questions specifically for the participant's version of the artefact. Here, the people seeing the score answered that they either stuck to one side or changed the alignment due to new information in the story. The people who did not see their score, on the other hand, were much more likely to stick to one alignment, with very few having a split or changing alignment. Looking at the individual answers for the person with the 6:4 split, as the outliers are important, shows that they chose that they changed their alignment due to new information while playing the version where they could see how their score changed. It is unclear in what direction their alignment changed at a certain point

in the story. They also say that they would have preferred not to see the score, as it was leading them to certain answers.

Taking the general answers into account for the question, it can be theorized that not seeing one's score actually made people feel like they were sticking to their guns in terms of alignment. The people seeing their score, when changing their opinion, instead felt like they switched alignment when choosing answers with another alignment. In simple terms; seeing the score and going to Advocate after previously having chosen an Ordinance response made them feel like they switched alignment. When not seeing the score change, going from Ordinance to Advocate without seeing those moral values assigned in the response made the player feel like they were still in the same alignment as they started even if there is some split in the actual scores. What this may point to is that if the player knows that they are being judged with a score, the responses are also considered more mechanically while choosing. This is opposed to when the player does not know their response has a value assigned to it, whatever they choose does not make them feel like they are "deviating" from an alignment, as they are simply choosing in accordance with their own one. The influence of seeing one's alignment is therefore not forced, but has power by suggestion.

Another observation made is the apparent disconnect between some players' views on The Devil as a role, and the main character of the game, which has been discussed in *5.1 The Study*. The relevant survey question can be seen in *figure 9*. The question asks how players made choices. I would like to also bring in the discussion from 5.1 The Study, about what makes something more or less moral when it comes to hurting others. This spawned from the questions about if the sinner should have treated her mother better (figure 11), whether revenge was her motivation (figure 12), if the responsibility was on her for the death of her mother (figure 13), and if the participants changed their mind on the sinner (figure 14). The responses across these suggest that the player's sense of ethos was engaged in an almost objective way in some cases. The participants felt some guilt or empathy in how they judged the sinner. The circumstances of her actions were by many seen as justified, even if the actions themselves were seen as "bad" actions. The individual responses still argue within the context of the game's rules, meaning from the perspective of being a judge. This means that the participants acted according to the logic of the game, rather than purely from their own real-life perspectives.

Ahn, Grizzard and Lee (2021) argue that justifying actions from the rules given to the players depend on their in-game role. Even though Ahn, Grizzard and Lee discuss this specifically in justifying immoral actions committed by a hero, the disconnect of The Devil as a role and as a character makes the view of that perspective character unclear. Some of these players could see The Devil as a sort of anti-hero, especially with rhetoric available in the game, where they

can be against the rule of God. Considering the prevalence of The Advocate alignment, most players acted as a Devil that sought justice in favour of a sinner they considered to be in Hell for the wrong reasons, not deserving that fate. Judgement can, in some cases, be seen as an immoral action. It is after all the act of imposing one's own morals on another. This was however not questioned, at least in what can be seen from the survey. The judgment was ostensibly seen as a morally acceptable action. In the discussion of The Devil it was noted that some may see them as antagonistic, or perhaps an anti-hero. An antagonistic character but a justified action (judgement) could answer the apprehension to see character and role as the same thing.

Either way, the unclear position the player holds in the game, while still viewing the situation quite logically, suggests high engagement. Consalvo, Bush, and Jong (2016) argue that high engagement in a game can result in hampered empathy in moral dilemmas. The participants of this study do mention empathetic reasons for why the sinner was/was not guilty, but some justified their actual decision on the rules of the context of the situation in the game. Others went the other way, saying why she was guilty or not, but then justifying their decision with a pathos argument. Unfortunately, there is not enough data to confirm if the version of the artefact affects this, but the artefact did in general engage both ethos and pathos in the player's decision. A morally grey character, like The Devil, may in fact be beneficial for this as it urges the player to consider both their own pathos/ethos, in addition to the game role's ethos. It also suggests that the participants were engaged with the dual-process theory (Greene, Nystrom, Engell, Darley, & Cohen, 2004). They make snap judgements based on their first impression of the situation, but is urged to consider the dilemma through cognitive judgement. This may be more difficult with a hero figure, as this type of character can be assumed to be just in their actions either way.

The second thing examined is what system was more engaging. Engagement in the survey is defined by reading the text without skimming, considering every option, and feeling invested in the story. For the group seeing the score, no one felt like they were necessarily pressured into sticking to one alignment. Note however, it can be seen that there have been mentions by this group that changing alignment was felt, whereas the other group did not feel like they changed an alignment when they according to their score did so. However, the consensus was that they did not believe that they were more engaged because of this system, and most also said that they would have preferred not to see the score. Some ethical cognitive dissonance is present here (Formosa 2022), meaning that the player feels like the mechanics of morality in the game works against their ability to engage with the actual moral dilemmas. The system did of course have positives as well, as some people felt like the score helped them make up their minds quicker than they think would have been the case without it. With these things in mind, it would seem that the participants believe that the positives don't outweigh the negatives, as

there is an acknowledgement of it being helpful, but they would like to play the game without seeing their score change. Perhaps this is because a player wishes to be challenged, whether that be in skill or in this case, moral judgement, but this is currently conjecture.

For the other version, the majority of participants said that not seeing the score made them engage more with the game (repeating the definition of engagement, it is to read the text without skimming, considering every option, and feeling invested in the story). Most also felt like they found the story more interesting than they believe they would have if they saw the score, as they weren't told what to think. They were also listening to other characters' opinions, which shows engagement. Only one person thought it made it difficult to make choices, or at least that such was a negative facet of the game. Perhaps most telling here is that no one in this group would have preferred to see the score change as they played. The conclusion for this group is quite clear; going in blind as to the values assigned to their choices make people feel more engaged with the material and the moral dilemmas in it.

In summary, the pool of participants was found to be mostly young atheists, with only a few religious people. Most of the participants chose more often in favour of the Advocate alignment, the alignment representing morality more forgiving of the character on trial for alleged "sinful" actions. The participants generally stuck to one alignment, with some outliers that were more split. Participants generally seemed to be engaged with both their own pathos/ethos, as well as the ethos given to them by the rules of the game. In the questions about the specific version of the artefact, ergo, either for players of the version showing alignment score changes in choices or not, they seemed to enjoy the version not showing them more. The one that showed them these changes made participants feel like they changed/broke their alignment if they switched the type of response later on, and felt less engaged with the material. Many of this group would have preferred the other version. The other group, not seeing score changes, felt more engaged with the story, characters, and the moral dilemma. None of this group wished they played the other version.

5.3 Conclusions

The questions for this thesis are whether seeing a moral alignment score or not make players feel more engaged with the game they are playing, and what system helps the player tackle a game's complex moral questions the best. From the results of the survey it can be seen that not seeing how a moral alignment score changes seem to be preferred in both cases. The conclusion from the results paints the picture that not seeing score changes also makes the player feel that they are freer to make their own choices. The lack of an external source that categorizes morality in their stead also made them feel like their choices were always in line with their own

moral judgement, while the smallest deviation in the version where the participants could see this instead made the player feel like they "switched". In the literature review, moral focus is mentioned as one of the facets of morality (Formosa, 2016). What this entails is having a moral identity, having consistency in what one chooses before moral dilemmas. As Formosa says, it is mainly an internal process and should not need external output to keep existing. A part of the internal process however, is that one can view oneself from the perspective of others. Moral meters are a mechanical version of this that is more or less forced onto the player. What it does is say "you made these choices, so this is the kind of person you are". As said, this can be beneficial in a game where morality is a game mechanic where the player benefits from certain scores, so that the player knows what they should be doing to go in the "right" direction. This can be particularly useful in a game with any factions that judge the player on their actions. The biggest strength of morality meters is the clear feedback. What the survey result shows however, is that it isn't necessarily enjoyable or engaging, at least in a game where morality isn't directly tied to rewards. This is likely because of the topic of moral focus. This is something that the participants seemed to like to establish themselves. The moral prompts are simply a tool for self-reflection. In retrospect, it may have been beneficial to include a question about the dynamic feedback present in the artefact, where characters both respond to the player's choices directly, and also bring them up later on. This question may be stated as "were the reactions of characters enjoyable and clear in respect to your choices?", or in the same vein. This would compare against the participants' opinions on the scoring system.

The participants were generally quite engaged with the moral material of the artefact, regardless of the version they played. They put thought into the issues. There may have been some suggestion in the version where there was a score visible, but the participants tried to play reflectively. The score was likely superfluous however, as it didn't add much to the experience except as a way to measure morality based on the values assigned to choices. In this game, where the main focus is complex moral issues, the score only hampered the intentions of the story. This could be viewed as an unfair disadvantage. It is however based in the issue of morality meters. An argument against them is that they can be almost pointless for many games as there are systems that lend themselves better to the type of game they are in, namely RPGs (Formosa, 2022). Most people play this type of game to personally affect the world and story. Dynamic feedback is better for this as the player won't feel like the game world changes due to numerical numbers, but due to their actions. With the morality scoring system, there's a bigger motivation to fill up on points than choosing earnestly, going against the point of a game like *Mass Effect* where you're meant to shape the galaxy's fate.

What Consalvo (2016) wrote about how engagement with a game lowering empathy was complex in this study. As established, the participants generally felt engaged. Many of the write-in questions go into the empathy they felt for the character on trial in the game. Despite

this, for a few of the participants their values played a bigger role than their empathy, deciding that even though the sinner's actions were justified, their role and convictions led them to choose against the character's innocence. They were also playing as a morally grey character (Ahn, Grizzard & Lee, 2021), which means that they had a strong moral focus as they couldn't justify their choices by saying that their avatar's choices are always good. They can therefore look past their empathy to choose what they thought to be the most moral. The reason so many felt both empathy and chose in favour for the sinner despite being engaged with the game is likely for the simple reason that their feelings and convictions aligned. These participants, the majority of the pool of participants, therefore did not have the same conflict of interest. However, regardless of what type of choices the participants made, there seem to have been thought put into the choice, suggesting that the dual-process model played a part in this thought process (Greene, Nystrom, Engell, Darley, & Cohen, 2004). This means that regardless of initial reaction to the situation, they still regarded the dilemma through their cognitive processes. They came to a conclusion based on reasoning in collaboration with their moral focus.

All in all, the answers to the thesis questions suggests that for the most beneficial experience for both the player and the creator of the material, not assigning value to moral choices and showing that to the player is the best course of action. The player will feel like they have more agency in the game, and will engage more with the narrative that the developer has worked hard on. They would also create a stronger moral identity in the game, as suggested by this study, while also being more receptive to questioning their own actions. Of course, there are use cases for other types of presenting morality to the player. This could open up to more criticism, and should be used with good reason. Consider first if there is a type of morality system that is generally more liked by audiences. If using these would be detrimental to the vision of the game, it is of course much better to be honest and risk something that may be less liked. It can be noted that none of the reading material used for this thesis was as negative towards morality systems other than morality meters, likely as this is the one that is the focus of discourse, having garnered a bad reputation for being used incorrectly.

6 Concluding Remarks

6.1 Summary

With the discourse surrounding morality in video games, and in particular how this is measured and presented to the player, this study aimed to investigate two of the commonly used methods. These were the morality meter type of moral score, where it is more or less evident to the player how this score changes as they play. The other method is more covert, tracking the player's choices, sometimes still assigning a value to them, but not making it oblique to the player as they play.

Two artefacts were created. These used the same story, with the same choices, valued in terms of a score in two alignments. The values were the same in both versions. One version told the player how their score changed according to their choices, while the other did not. The study also asked the participants about their opinions of the content in the artefact to establish if there was a relation to how they play a game, and how their morality works. Generally, the participants were forgiving of the central character in the game, despite what may be considered wrongdoings. No relationship between this aspect and the version of moral score was discerned. However, most participants preferred the version where the score was not apparent to them while playing, where the people seeing it said that they would have preferred not to, and the people not seeing the score change were happy with the system. The participants also felt more engaged with the story if they could not see the score, whereas the other group believe that they would also feel more engaged if they did not see the score.

6.2 Discussion

The major concern with this study was that the groups were not symmetrical, where the group not seeing how their score changed was larger than the other. A more even grouping may have given room for more outliers, or types of players, to present their opinions. This would have given more room for comparison of the two groups. A better distribution may have given a clearer overview of how enjoyable the experience was for the groups. As it now stands, one group gets more representation, meaning that the group seeing the score may have been unfairly judged. A larger pool of participants in this group may have given a more positive impression of the visible score. Alternatively, it could also have given a stronger base to dismiss the system if more participants had a negative view of it.

Additionally, a major component of the artefact was theist and moral philosophy. Having a larger number of religious participants would likely have given more varied results in terms of what choices the participants made. Not only would there be great differences between religious and atheist respondents, but between religions as well. Cohen and Rozin (2001) shows this, explaining how the concept of sin is fundamentally different in Judaism and Christianity.

What the study may be useful for is that almost every participant believed that not being constantly reminded of their moral tracking would result in a higher degree of engagement with the material. This can be important for developers to keep in mind. Whether or not engagement *is* higher, perception is a powerful tool to tell good stories and reaching a wider audience. A decent story in video game format can be thought to be better than a great one if the decent one makes the player feel like they have agency within it. Furthermore, the study gives the impression that not seeing moral score changes was preferred over seeing them. From the responses from the participants, it is clear that being told that their choices already have values assigned to them was taken negatively. As discussed in Conclusions, morality meters, or moral systems apparent to the player, are stigmatized within the gaming community. This study does confirm this, while also establishing that these concerns can be valid. The study recommends that for narrative games, including these systems should be gauged to see if it is actually the best course of action for the particular game the developer is creating. The reason is that the player will in many cases go into it already having a negative image of these systems from past experiences, and therefore be extra critical towards it.

6.3 Future Work

The topic would benefit from further studying. Since the popularity of especially *Mass Effect* and *Fallout*, many games tout the sentence “Your choices will change the story!”. The choice aspect, or “play YOUR way”, is a massive selling point for narrative-focused games, and for good reason. People who enjoy narrative games often play games because they wish to have agency in the story. Without it, they may as well watch a movie instead. Due to the importance of getting things right, it is particularly important to give game developers the correct assortment of tools to create stories that hold up to the expectations and hopes of the audience.

Many aspects affect what makes the right morality system for a game. How linear is it? The resources needed to create a linear game and one with branching paths will differ greatly. Not only would it take more time and money to create something that branches, but the construction of the game would also be different as so many moving paths must fit together. A morality meter could in fact be more beneficial in a case like this, as characters and events

could react to the player character as a whole instead of lesser individual choices. Another angle could be, how many moral alignments/factions would there be (Pluralsight, 2015)? Similarly to linearity, having ten factions could be too convoluted to keep track of if one was to use a dynamic system. This would be particularly interesting to study if these factions also intersect where some are allies or enemies. Furthermore, will the player make choices at pivotal moments, or is it the small actions within gameplay that count (Formosa 2013, 2022)? This could compare morality systems that are narrative vs mechanical, something that was outside the scope of this study. An angle to this question would be to see what is more enjoyable, or even categorize the ideal system for different types of games. One may also look into what caters best to reactive or reflective players (Formosa 2022). This is a topic that often come up in studies and literature about video games. Often, a game will try to appeal to both types. Another study may go away from morality and instead focus on what group is more profitable, or if spreading yourself too thin harms the experience for both groups.

This study in particular has looked at engagement and player preference between meter-based vs covert morality systems in short and linear narrative games. For a study to build upon this one, it would be greatly beneficial to categorize more genres of games to give developers clear blueprints and guidelines for what system is most effective for the type of game or story they want to create. It can be daunting to choose something that there is uncertainty in whether or not it will work the way they wanted it to.

Looking at the issues found within this particular study, there may be some things to think about in further studies. The number of participants was not equally distributed according to the number of artefacts used. The number of participants could also be higher to get more raw data, and increase the responses that are outliers to the norms. The participants were also mostly gathered from Sweden and Finland. Including theist morality is absolutely not important for the topic of moral alignments, but if this is something that is interesting in further studies, sourcing participants from countries where atheism is not as prevalent would be very beneficial for more extensive data in this area of research. The theist morality in this study was mostly seen through the perspective of atheist views. This could explain the overwhelming preference for the moral alignment that opposes the moral alignment informed by Christian morality. These results could look very different if, for example, the participants were from Italy or Armenia where a majority of people are Christians. The way a culture's morality functions is highly influenced by religion, as seen in Jun's study (2005). The video game format could be used in studies of religious differences as seen in Cohen and Rozin's study (2001). Choices in studies pertaining to theology could include ethics from different ways of thinking to see consistency in religious groups, or the prevalence of religious doctrine in secular society. Morality in games is a very expansive area of study, and can even form the basis of studies outside the game industry.

References

Ahn, C., Grizzard, M., Lee, S. (2021). How do video games elicit guilt in players? Linking character morality to guilt through a mediation analysis. *Frontier in Psychology,* 12. DOI: https://doi.org/10.3389/fpsyg.2021.666518

Bateman, S., Birk, M., Pradantyo, R. (2021). How the visual design of video game antagonists affects perception of morality, *Frontiers in computer science*. *10.3389/fcomp.2021.531713*.

Bond, B. (2021). The development and influence of parasocial relationships with television characters: A longitudinal experimental test of prejudice reduction through parasocial contact. *Communication Research*, 48(4): 573-593. DOI: 10.1177/0093650219900632

Cohen, A., Rozin, P. (2001). Religion and the morality of mentality. *Journal of personality and social psychology*, 81(4): 697-710. DOI: 10.1037/0022-3514.81.4.697.

Consalvo, M. Busch, T. Jong, C. (2016). Playing a Better Me: How Players Rehearse Their Ethos via Moral Choices. *Games and culture*, 14(3): 216-235. DOI:10.1177/1555412016677449

Creswell, J. (2009). *Research design: Qualitative, quantative, and mixed method approaches*. Thousand Oaks, CA: SAGE Publications Inc.

English standard version Bible. (2016). Crossway. https://www.biblegateway.com/versions/English-Standard-Version-ESV-Bible/#vinfo

Fallout 3. (2008). Bethesda Softworks. [PC, Xbox 360, Playstation 3]. Available at: https://store.steampowered.com/app/22300/Fallout_3/

Formosa, P., Ryan, M., Howarth, S., Jane Messer, J., McEwan, M. (2022). Morality Meters and Their Impacts on Moral Choices in Videogames: A Qualitative Study. *Games and Culture*, 17(1): 89-121. DOI: 10.1177/15554120211017040

Formosa, P., Ryan, M., Staines, D. (2016). Papers, please and the systemic approach to engaging ethical expertise in videogames. *Ethics and Information Technology*, 18(3): 211-225. DOI: 10.1007/s10676-016-9407-z

Freeman, T. (2020). Do Jews believe in heaven? *Chabad*. https://www.chabad.org/library/article_cdo/aid/4848230/jewish/Heaven-and-the-Afterlife.htm [17/02/2022].

Greene, J. (2009). Dual-process morality and the personal/impersonal distinction: A reply to McGuire, Langdon, Coltheart, and Mackenzie. *Journal of experimental social psychology*, 45(3): 581-584. DOI: 10.1016/j.jesp.2009.01.003

Greene, J., Nystrom, L., Engell, A., Darley, J., & Cohen, J. (2004). The neural bases of cognitive conflict and control in moral judgement. *Neuron,* 44: DOI: 10.1016/j.neuron.2004.09.027

Hodge, S., Taylor, J., McAlenery, J. (2019). (A)morally demanding game? An exploration of moral decision-making in a purpose-made video game. *Media and communication*, 7(4): 213-225.

Infamous. (2009). Sony Computer Entertainment. [Plastation 3]. Not available to purchase online.

Jun, S. (2005). Moral differences between protestants and atheists, and the effects of religion on the morality of protestants. *Korus*, DOI: 10.1109/KORUS.2005.1507966

McGuire, J., Langdon, R., Coltheart, M., Mackenzie, C. (2009). A reanalysis of the personal/impersonal distinction in moral psychology research. *Journal of Experimental Social Psychology,* 45(3): 577-580. DOI: https://doi.org/10.1016/j.jesp.2009.01.002

Papers, Please. (2013). 3909 LLC. [PC, Mobile]. Available on the internet: https://papersplea.se/

Phillips, P., Aaron, B., Phillip, J. (2013). *Survey Basics*. Alexandria, VA: American Society for Training and Development.

Pluralsight. (2015). 5 Steps to Responsibly Designing Morally Difficult Characters. *Pluralsight*. [Blog] March 4. https://www.pluralsight.com/blog/film-games/5-steps-responsibly-designing-morally-difficult-characters?exp=2 [30/01/2022]

Shin Megami Tensei. (1992). Atlus. [Super Famicom]. Not available to purchase online.

Sicart, M. (2013). *Beyond choices: The design of ethical gameplay*. Cambridge, MA: The MIT Press.

7 Appendix

7.1 Appendix A

Survey questions

Morality Survey: Intro

This is the survey part of the morality in games study. This should be done only after playing either The Devil's Throne Prologue Ver.A, or The Devil's Throne Prologue Ver.B. Only one should be played. These can be found on here:
Ver.A: https://kylar-pink-narcissus.itch.io/the-devils-throne-vera
Ver.B: https://kylar-pink-narcissus.itch.io/the-devils-throne

* Required

1. What age are you? You may skip this question.

   Mark only one oval.

   - ◯ -18
   - ◯ 18-25
   - ◯ 26-35
   - ◯ 35-65
   - ◯ 65+

2. Would you say that you identify as a follower of a religion? *

   Check all that apply.

   - ☐ No, agnostic
   - ☐ No, atheist
   - ☐ Yes, Christian
   - ☐ Yes, Jewish
   - ☐ Yes, Muslim
   - ☐ Yes, but not an Abrahamic religion
   - ☐ I do not know
   - ☐ Other: _____

3. What type of games, if any, do you usually play. Up to 3. You may skip this question.

Check all that apply.

- [ ] I do not play video games at all/very little
- [ ] Puzzle Games (Candy Crush, Tetris)
- [ ] Story Driven Adventure (Life is Strange, Night in the Woods, Oxenfree)
- [ ] Visual Novels (Doki Doki Literature Club, Dream Daddy)
- [ ] First-Person Action Shooters (Call of Duty, Battlefield)
- [ ] Third-Person Action Shooters (Gears of War)
- [ ] Competitive Shooters (Fortnite, Apex Legends)
- [ ] RPGs (role-playing games) (Skyrim, Mass Effect, Undertale)
- [ ] Simulation Games (Minecraft, Stardew Valley)
- [ ] Adventure Games (Uncharted, Zelda: Breath of The Wild)
- [ ] Action Game (God of War, Assassin's Creed)
- [ ] Fighters (Tekken, Dragon Ball)
- [ ] Horror (Silent Hill, Outlast)
- [ ] Walking Sims (Dear Esther, The Stanley Parable)
- [ ] Strategy Game (StarCraft, Crusader Kings)
- [ ] MMO (World of Warcraft, Final Fantasy XIV)
- [ ] Sports Games (FIFA, NHL)
- [ ] Racing Games (Forza, Mario Kart)
- [ ] Other: _____

4. What were the scores you got in the end of the game that's part of the study? (Please clarify with "Advocate" and "Ordinance" plus a number. Example; Advocate: 2 Ordinance: 8.). If you do not know, leave blank.

_____

Morality Survey: About Choices

5. Do you think The Sinner should have treated her mother better? Feel free to write in-depth in the Other choice. *

Check all that apply.

- [ ] Yes
- [ ] No
- [ ] Not sure
- [ ] Other: _____

6. Was The Sinner motivated by revenge? Feel free to write in-depth in the Other choice. *

Check all that apply.

- [ ] Yes
- [ ] No
- [ ] Not sure
- [ ] Other: _____

7. Is The Sinner responsible for her mother's death? Feel free to write in-depth in the Other choice. *

Check all that apply.

- [ ] Yes
- [ ] No
- [ ] Not sure
- [ ] Other: _____

8. What made you pick the choices you did? *

Check all that apply.

- [ ] I chose what I would in real life
- [ ] I chose according to what made sense for the main character
- [ ] I chose what I would in the "role" as The Devil
- [ ] I chose what I thought was good/right
- [ ] I chose what I thought was evil/wrong
- [ ] I chose to maximize my Advocate/Ordinance score
- [ ] I chose so that the companion I preferred would like me
- [ ] Other: _____

9. When you play games with choices that deal with dilemmas of right and wrong, or some variation of that (Mass Effect, Undertale, Bioshock), how do you engage with the game? *

Mark only one oval.

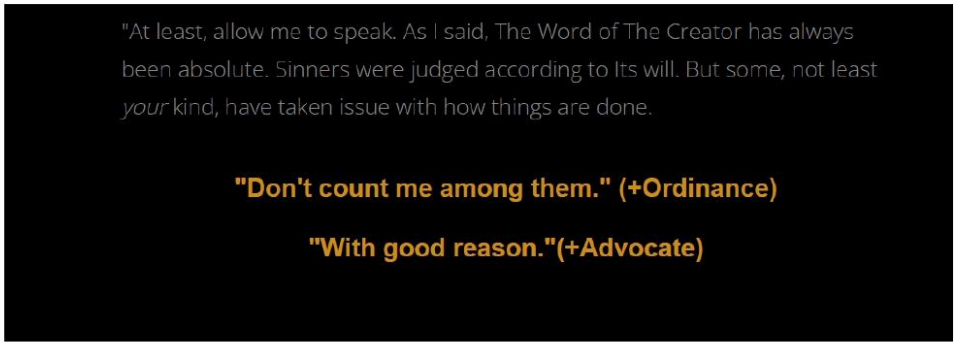
- ◯ I try to get the best rewards regardless of if some choices do not align with my ideals, but would not look up how to do it.
- ◯ I try to get the best rewards regardless of if some choices do not align with my ideals, and may look up how to do it.
- ◯ I do not play for reward and take things as they come.
- ◯ I don't play games
- ◯ Other: _____

10. Do you feel like your opinion on The Sinner changed through the story? *

Mark only one oval.

- ◯ Yes, I sympathized more
- ◯ Yes, I sympathized less
- ◯ No change, always negative
- ◯ No change, always positive
- ◯ Other: _____

For question below.



11. In the game, did some of the choices tell you that picking them would give a point in a certain category? See picture above as an example. *

Mark only one oval.

- ◯ I did not see that *Skip to question 15*
- ◯ I did see those *Skip to question 12*
- ◯ I do not know

Questions for Ver A

12. If you were shown how your score was affected by your choices, how did you experience this?

Check all that apply.

- ☐ I would have preferred to not see how my score was affected.
- ☐ I felt like it helped me stay on the course I wanted.
- ☐ I felt like I was pressured into sticking to the side that I started with.
- ☐ It helped me make up my mind.
- ☐ I ignored it.
- ☐ It made me question myself/my morality.
- ☐ Other: _____

13. Did seeing how a score changes make you engage more with the narrative? (Ex: reading the text closely without skimming, considering every option, feeling invested).

Check all that apply.

- [ ] I think so
- [ ] I do not think so
- [ ] Other: _____

14. Did you stick to mostly one alignment?

Check all that apply.

- [ ] Yes, I was almost exclusively picking one side
- [ ] No, it was pretty much 50/50
- [ ] I changed my alignment due to new information in the story

Questions for Ver B

15. If you did not see how choices affected a score, how did you experience this?

Check all that apply.

- [ ] I would have preferred to see how my score was affected.
- [ ] The absence of a visible score made me question myself/my morality.
- [ ] The absence of a visible score made the story more interesting as I was not told what to think.
- [ ] It was difficult to make choices as I did not know how they would affect the story.
- [ ] It made me listen to other characters' opinions.
- [ ] It made me ignore the other characters' opinions.
- [ ] Other: _____

16. Did the lack of a visible score make you engage more with the narrative? (Ex: reading the text closely without skimming, considering every option, feeling invested).

Check all that apply.

- [ ] I think so
- [ ] I do not think so
- [ ] Other: _____

17. Did you always pick either mostly The Advocate's or the Ordinance's side, or were your choices more fluid?

Check all that apply.

- [ ] I picked mostly one side
- [ ] My choices were mixed
- [ ] I started out with one side, but shifted due to new information in the story

This content is neither created nor endorsed by Google.

Google Forms

7.2 Appendix B

Links to playable versions of the Artefacts.

Artefact version A, or version with visible alignment score changes in choices: https://kylar-pink-narcissus.itch.io/the-devils-throne-vera

Artefact version B, or version without visible alignment score changes in choices: https://kylar-pink-narcissus.itch.io/the-devils-throne