UNIVERSITY *of*
TASMANIA
AUSTRALIA

PERSONALITY ASSESSMENT THROUGH THE USE OF VIDEO GAMES

By

Adam Zulkifly

(BComp, Hons)

Submitted in fulfilment of the requirements for the degree of

Doctor of Philosophy

University of Tasmania

July 2019

Declaration of Originality

This thesis contains no material which has been accepted for a degree or diploma by the University or any other institution, except by way of background information and duly acknowledged in the thesis, and to the best of my knowledge and belief no material previously published or written by another person except where due acknowledgement is made in the text of the thesis, nor does the thesis contain any material that infringes copyright.

Authority of Access

This thesis may be made available for loan and limited copying and communication in accordance with the Copyright Act 1968.

Statement of Ethical Conduct

The research associated with this thesis abides by the international and Australian codes on human and animal experimentation, the guidelines by the Australian Government's Office of the Gene Technology Regulator and the rulings of the Safety, Ethics and Institutional Biosafety Committees of the University.

Signed: Adam Zulkifly

Date: July 3, 2019

Abstract

Research into serious games has increased over the recent years, but it remains a field that has vast potential applications when paired with other fields. This project seeks to situate a serious game within the area of personality psychology research with the aim of developing a tool that aids the identification and elicitation of personality information from individuals. This thesis describes the exploratory process of adapting an established theoretical framework into the design process of a serious game and the challenges that arise from such an endeavour.

A serious game is described, developed, and preliminarily tested in order to evaluate the design and implementation process of the video game tool. The research thus highlights several lessons learned during this process which will be able to be applied to future serious game development in order to create better tools that are more grounded in theory.

Acknowledgements

First and foremost, I have to thank the people who made this possible at all: Dr. Kristy de Salas and Dr. Ian Lewis. I shudder to even think about any of this process without your guidance. To say that your advice, experience, and steadfastness is invaluable is an understatement. I'm very far from eloquent at this point, but I hope I can somehow convey how much you did for me and for the project. Thank you both so much for not just putting up with me, but supporting me through all the headaches I caused that you most certainly didn't deserve.

Thank you to Professor Ken Kirkby and Dr. Allison Matthews. I don't want to imagine what it would have been like to attempt this without your knowledge. Thank you so much for putting up with what must have been amateurishly laughable attempts at learning something so complex in such a short time span. Thank you for your patience in explaining the intricacies of a field I walked into wholly unprepared for.

Thank you to the fellow crazy people in the GRG: Anthony, Aran, Ben, and Lindsay. Knowing that someone else is going through the same things you are and the solidarity that provides is really nice and helps remind me that everything is totally manageable. Just have to buckle down and do it.

To my friends. I consider myself so very fortunate that it doesn't make sense to try and list names here because of how many there are. How very fortunate I am that there are so many people around me who offer encouragement and understanding through this process. You may never know how much it helped just to have someone there to talk with, play games with, or just generally be idiots with and how all of that helped preserve my sanity (even if I didn't get to keep completely sane).

Lastly, but certainly not least, thank you to my family. Sometimes the unflinching belief in me is scary, but it is also very comforting to have you guys at my back. Thank you mum, dad, and Tania. I owe more than I probably even realize to you guys.

Contents

List of Figures

List of Tables

1. Introduction

This thesis discusses the creation and testing of a video game tool that is designed to assess the personality of the player. This includes situating the entire project within academia and the research into serious games, elaborating on the area of personality psychology as well as the need for such a tool, and then a full description of the process of creating that tool. This is followed by the testing of the designed video game tool, its results, and a reflection on the entire project.

1.1 Chapter Overview

Chapter 2 provides an introduction to video games and more specifically serious games. It highlights prominent examples of serious games and introduces the area of personality. Following this, the opportunity that exists in the overlap between serious games and personality is articulated and a research question is presented that can address this.

Chapter 3 discusses a methodology to answer the research question by systematically answering sub-components of the larger question. The process of deriving a game design from the theory of personality is presented. A description is provided here of the video game tool that was implemented from that design as well as the experimental procedure used to test it.

Chapter 4 presents the results from testing the video game tool.

Chapter 5 discusses the results shown in Chapter 4 as well as reflects on the lessons learned through the entire process in order to answer the research question.

Chapter 6 summarizes the work that has been done through this project.

2. Literature Review

This chapter will discuss the relevant literature background to this project in order to establish the opportunity space that this project seeks to address. This discussion begins with a description of the increasing impact of video games and how the medium has expanded beyond strictly entertainment purposes and continues on into the area of personality research and the opportunity for video games to be used as a tool for personality data elicitation.

2.1 Literature Review: Video Games

Video games are currently an extremely prolific and ubiquitous medium that is consistently growing and reaching wider audiences (Entertainment Software Association 2017; IGEA 2017). As of 2016, the video games industry grossed 24.5 billion USD from the purchase of video game software and hardware (e.g. video game consoles such as the Sony PlayStation 4, Microsoft Xbox One, and Nintendo WiiU) in the United States alone (Entertainment Software Association 2017), and 2.958 billion AUD in Australia (IGEA 2017).

These figures in the commercial games industry are mirrored by a similar increase in interest in the academic community, with Universities offering video game degrees and majors[1] and numerous conferences and journals in the area[2].

Research on video games has not only investigated how to improve the hardware and software used by commercial video games which traditionally have the purpose of entertainment (Egenfeldt-Nielsen et al. 2009; Robertson 2012), but also explores the concept of using video games in non-entertainment and non-gaming contexts which are called *serious games* (Rooney 2012; Liu et al. 2013). As this chapter will show through examples, serious games have had an impact in a number of areas, including but certainly not limited to education (Ke 2012; Emam&Mostafa 2012), training (Nieborg 2004; United States Army 2002;

[1] University of New York - https://gamecenter.nyu.edu/academics/courses/
University of the Creative Arts - https://www.uca.ac.uk/study/courses/bsc-computer-games-technology/
University of Swinburne - https://www.swinburne.edu.au/study/find-a-course/games-animation/games-development/
University of Utah - https://eae.utah.edu/files/2018/01/BSG-Jan-2018.pdf
RMIT - https://eae.utah.edu/files/2018/01/BSG-Jan-2018.pdf

[2] ACM Computers in Entertainment - https://cie.acm.org/
Games and Culture - http://journals.sagepub.com/home/gac
Journal of Games Criticism - http://gamescriticism.org/

Fitz-Walter et al. 2011; Fitz-Walter et al. 2013), therapy (McGonigal 2012), health (Edgerton 2009), improving web services (Von Ahn&Dabbish 2004), and scientific discovery (Khatib, Cooper, et al. 2011; University of Washington 2008).

2.1.1 Video Games in Academia

Although video games have been the subject of research for some time, 2001 saw the birth of the first peer-reviewed academic journal that was devoted entirely to the research of computer games – titled Game Studies (Aarseth 2001). The year 2001 also saw the first year of the first international scholarly conference on computer games: *Computer Games & Digital Textualities* (DDCA 2001).

Since 2001, video game studies continue to be primarily multidisciplinary, with video game research overlapping into areas such as psychology, sociology, anthropology, and of course straight up technology standpoints (among many others) (Sheng-Yi et al. 2012; Yang et al. 2014; Jaakko 2012; Lofgren & Fefferman 2007; Van Lankveld, Schreurs, et al. 2011).

Research in the area of video games has propelled technology forward both in terms of software and hardware, but more interestingly is the application of the results from games and games research for non-entertainment purposes – serious games which are described in the following section.

2.1.1.1 Serious Games

Serious games are video games which are designed to deliver more than just entertainment to the player (Charsky 2010).

One of the earliest concepts of a serious game was defined by Clark Abt in his book titled *Serious Games* (Abt 1970) which focused mostly on board and card games and asserted that serious games:

> *"have an explicit and carefully thought-out educational purpose and are not intended to be played primarily for amusement"* (p. 9)

This definition specifically highlights an educational purpose to serious games as well as the intention of play to be something other than solely entertainment driven. As time has progressed and more focus shifted to the area and potential of serious games, multiple

definitions have surfaced (Marsh 2011). Zyda (2005) provides a definition for serious games as:

> *"a mental contest, played with a computer in accordance with specific rules, that uses entertainment to further government or corporate training, education, health, public policy, and strategic communication objectives."* (p. 26)

while (Michael & Chen 2006) define serious games as:

> *"games that do not have entertainment, enjoyment, or fun as their primary purpose."* (p. 21)

and (Susi et al. 2007) offers the definition of a serious game as:

> *"games that engage the user, and contribute to the achievement of a defined purpose other than pure entertainment (whether or not the user is consciously aware of it)."* (p. 5)

A concept that is repeated between the definitions is that of entertainment and achieving another purpose. The differences in opinion of what that other purpose may be is an indication of the diverse perspectives on serious games (Marsh 2011). Although there is disagreement on the importance of entertainment within a serious game, (Marsh 2011) goes further to say that:

> *"in general, it's not important whether or not entertainment or purpose is of primary importance, but the crucial issue is that the purpose is to some degree successful."* (p. 62)

With that said, this thesis asserts that for a serious game, the entertainment component is critical to a serious game. However, an equal measure of effort must be put into the overarching purpose of the serious game in order to achieve a successful serious game.

Perhaps the best way to convey the concept of a serious game is by highlighting notable examples in the following sections.

2.1.1.1.1 America's Army

Serious games are able to provide an insight into areas of expertise that an ordinary person would normally be unable to participate in. The *America's Army* video game is an example that has been funded and developed by the United States Army (United States Army 2002). America's Army was designed with the purpose of providing an engaging, informative, and

entertaining soldier experience through a virtual medium (McLeroy 2008). The game is primarily played from a first-person perspective, colloquially known as a first-person shooter (FPS) game as can be seen in Figure 1(Inc. 2002). The player begins in basic training missions which teaches them the game's controls as well as textbook military principles which includes assault rifle target practice (shooting in-game), obstacle course traversal (movement controls), and special U.S. weapons handling (such as how to prime or 'cook' a grenade) (Inc. 2002).

Figure 1 – America's Army basic training

After completing the mandatory basic training section, the player is allowed to pursue mode advanced training such as airborne school, medic training or advanced marksmanship where the game places emphasis on real physics, ballistics, fire rates, and assorted other problems and challenges with that training (Inc. 2002). Once the player is done with the instructional sections, they are able to partake in full-scale combat operations simulations where teams comprised of dozens of individuals pit themselves against each other in an assortment of environments and mission types that range from bridge assaults to fortified building raids to stealthy extraction missions such as shown in Figure 2 (Inc. 2002).

Figure 2 – America's Army combat operation mission

The game's emphasis on realism forces typical gameplay away from gung-ho heroism in favour of more tactical positioning, movement, and communication across a squad of players (Inc. 2002).

While entertainment was certainly also one of its goals, *America's Army* is also intended to be a strong recruitment tool for new soldiers – which has netted it criticism as being a propaganda device (Nieborg 2004; Delwiche 2007). In line with its goal of recruitment, the game is heavily coupled with the *Go Army* recruitment website.

Advances in game technology since its inception allowed *America's Army* to grow in terms of complexity to the point where it became useable as a training device that now sees use by other U.S. government departments such as the Secret Service (Zyda 2005). Anecdotal reports claim that recruits who struggled with the rifle range or obstacle course were able to pass the tests after playing the corresponding sections within the game (Zyda 2005).

In summary, the attention to finer real world details such as progressing the player steadily through every echelon of the U.S. Army and an emphasis on tactical combat scenarios allowed players to learn real-world Army tactics and habits while playing a video game.

2.1.1.1.2 Pharmacy Simulator

In a similar vein to the results of *America's Army* above, some serious games have been developed to provide training in ways which would not be practical or possible in the real world due to limitations such as safety, cost, or time. *Pharmacy Simulator* was designed to be a virtual environment where University level pharmacy students may practice responding to a wide range of scenarios (Bindoff et al. 2014).

The game features a computer-based simulation of a community pharmacy that is complete with a front desk, front of shop area, dispensary, dispensing computer, and telephone (Bindoff et al. 2014). Players control an avatar within the world from a first-person perspective and are given complete freedom to any of the aforementioned areas of the pharmacy (Bindoff et al. 2014). Administrators of the simulation program are able to write highly detailed and customizable scenarios for players to interact with and respond to as if they were working within a real life pharmacy (Bindoff et al. 2014). Players may be faced with scenarios that require them to elicit information from patients who walk into the pharmacy (as seen in Figure 3), follow proper dispensary procedures for retrieving medicine, or even simply being required to perform inventory checks on stock within the premises (Bindoff et al. 2014).

Figure 3 – Patient interaction in Pharmacy Simulator

Using a virtual simulation such as *Pharmacy Simulator* allows educators to place students into real world scenarios without any of the real world risks in order to put skills gained in theoretical lessons and discussions to some degree of practice as it is rarely practical nor affordable to have students in real world pharmacy placements all the time. Student groups reported enjoyment when using the simulation, though the authors note that further scenario development was required to obtain significant improvements in knowledge acquisition (Bindoff et al. 2014).

2.1.1.1.3 FoldIt

Serious games are able to leverage human problem solving capability and also can enable groups of people to compete or cooperate to solve the same problem. The game *Foldit* is a puzzle game that is distributed online where players manipulate protein structures to solve complicated protein structure prediction problems (Cooper et al. 2008).

Players are given a 3-dimensional (3D) representation of a protein structure and are scored based on how well they fulfill three criteria (Cooper et al. 2008). The first is to pack the protein to make the protein structure occupy less space by placing atoms within the structure to be as close together as possible (Cooper et al. 2008). The second criteria is to hide the hydrophobics – components of a protein structure marked orange in the game (see Figure 4) – as these should be surrounded by as many atoms as possible to minimize contact with water (Cooper et al. 2008). The third and final criteria is to clear clashes where two atoms are occupying the same space at the same time as this would be physically impossible in the real world (Cooper et al. 2008).

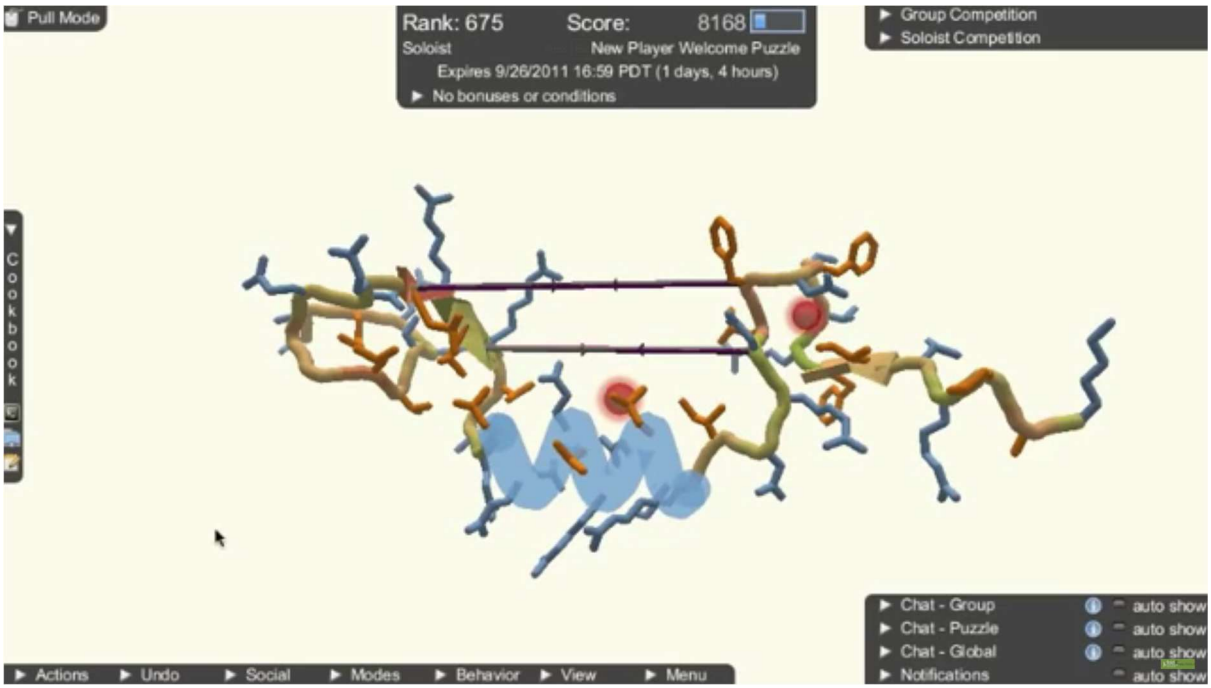

Figure 4 – Protein folding in Foldit

The website where *Foldit* is downloaded also offers education on the specific nature of proteins, why and how they fold, how the game helps the academic community, as well as a list of publications that have resulted from the game being played (Cooper et al. 2008). The website also offers a place for a community to develop around the game via contests, forums, and a wiki-database so that players will always be able to find a new challenge or share their accomplishments within the game (Cooper et al. 2008).

In summary, the *Foldit* game takes advantage of humans' abilities to solve 3D problems and uses information gained from players to improve algorithms used by computers to tackle those problems while offering players a way to compete with one another while simultaneously working towards a common goal (Cooper et al. 2011). The game has already proven to be successful as strategies formulated by players outperformed previously published methods and eventually demonstrated altogether new algorithms (Khatib et al. 2011).

2.1.1.1.4 ESP Game

It is also possible for serious games to use problems that the software itself is unable to solve and instead rely on humans to check and balance each other in a game format. The *ESP Game* pairs players and then shows both players identical images (as shown in Figure 5) and then

asks that they guess what the other player is typing (von Ahn & Dabbish 2004b). The likely result is for a player to describe the image in the hopes that their counterpart does the same, but in truth the game only cares if the textual input from both players match (von Ahn & Dabbish 2004b). Players are able to continuously make guesses until an 'agreement' is reached, but neither player can see the others' guesses (von Ahn & Dabbish 2004b).

Figure 5 – The ESP Game, showing an example of guesses by two players on a picture of a handbag

The game is further complicated by 'taboo' words which are generated for each image based on what previous player pairs had agreed upon for that image as can be seen in Figure 6 (von Ahn & Dabbish 2004b). The taboo words serve two roles: to make the game more challenging, and to increase the number of words that are returned for each image (von Ahn & Dabbish 2004b)

Figure 6 – The ESP game showing players taboo words based on previous guesses

13,630 players generated nearly 1.3 million labels within the first four months of release with promising results in terms of accuracy and labelling frequency (von Ahn & Dabbish 2004b). The rights to the game has since been obtained by Google to be used in conjunction with their *Image Search* service (von Ahn & Dabbish 2004a).

2.1.1.1.5 Quittr

Some serious games are designed with the goal of encouraging behaviour change in its players. *Quittr* is a serious game designed using motivational theory to encourage smokers to quit smoking (Bindoff et al. 2016).

The *Quittr* game aims to help users complete a 28 day no-smoking period (Bindoff et al. 2016). To that end, it features a framework that tracks their progress through the 28 day period as well as two minigames designed to fulfill a specific role in aiding a player quit smoking (Bindoff et al. 2016). The support framework tracks the player's progress through the 28 day period (as seen in Figure 7) as well as provide a range of helpful information and support which includes educational material such as information about various therapies, treatment plans, as well as information on cravings and coping strategies (Bindoff et al. 2016).

The two minigames (as seen in Figure 8) currently available with the *Quittr* game fulfill one of two functions: distraction or incentivization (Bindoff et al. 2016). The distraction minigame is designed to be played in a stand-alone 1- to 5- minute session in order to provide an effective distraction from the act of smoking (Bindoff et al. 2016). This game is designed to demand mental focus and the use of both hands, requiring the player to search for and tap particular objects in an increasingly cluttered environment under time pressure (Bindoff et al. 2016). The incentivization game is designed to provide tangible goals and rewards for the player over the 28-day quitting period via a resource management style game where the player is attempting to construct a city (Bindoff et al. 2016). The player invests resources into the town to grow it which in turn passively generates resources to be used to continue the town's expansion (Bindoff et al. 2016). The rate of resource gain is designed in such a way so as to ensure that the player will have a productive activity in the town every 1 to 2 hours (corresponding with when the player is likely to suffer from a cigarette craving) (Bindoff et al. 2016).

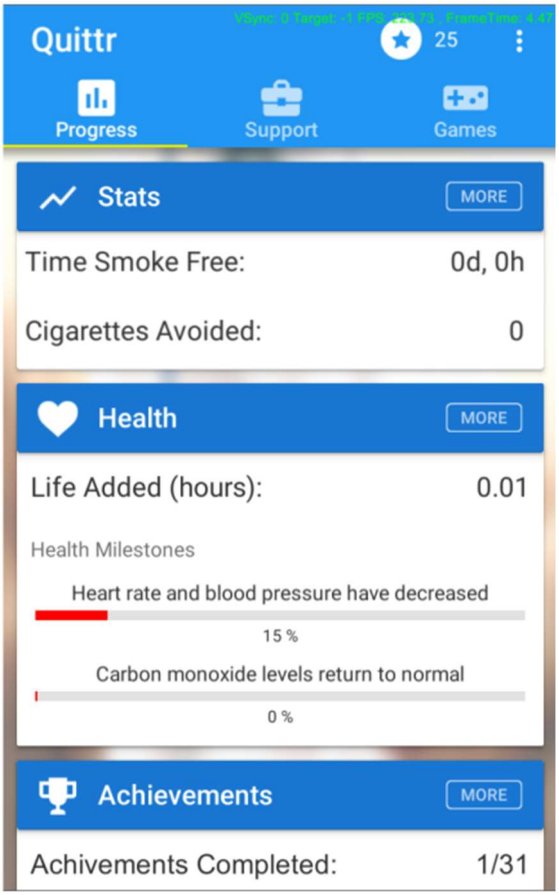

Figure 7 – Quittr tracking player progress over time

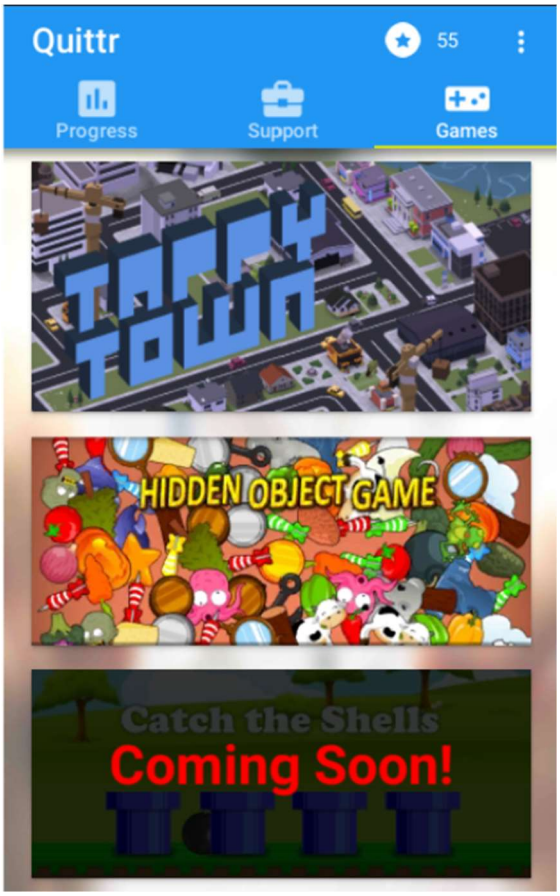

Figure 8 – Quittr's two minigames; Tappy Town and Hidden Object Game

The *Quittr* game is designed to be a comprehensive mobile application that is designed to aid users in quitting smoking which it accomplishes by using minigames as a means of providing both a distraction from smoking urges as well as a more tangible long term achievement through the use of video games that have been designed with motivational theory in mind (Bindoff et al. 2016).

2.1.1.1.6 Augmented Exercise Machines

While some serious games like *Quittr 2.1.1.1.5* are designed with the intent of completely alter an individual's behaviour, others have been developed to help individuals maintain the habit of beneficial behaviours. A 2016 study described a series of traditional exercise machines that were augmented to be equipped with video games that responded to the physical exertion on the machines (Geelan et al. 2016).

The study featured two traditional exercise machines: an articulated exercise bicycle and a rowing machine as can be found in a typical gym that were augmented to act as input for two commercial games: *Spin or Die* (Play 2016) and *Wii Sports Resort: Canoeing* (Nintendo 2009)(see

Figure 9 and Figure 10). The exercise bicycle is connected to the video game such that as the user is able to control the player character on screen using buttons near the handles of the exercise machine while pedaling faster increases the speed of the character on screen (Geelan et al. 2016). The rowing machine is connected to the video game by attaching a *Wii remote* from the Nintendo *Wii* console to the rowing oar of the exercise machine so that the player controls the character on screen by physically using the rowing exercise machine (Geelan et al. 2016).

Figure 9 – Modified cycling machine and game

Figure 10 – Modified rowing machine and game

The study found that although the augmented exercise machines didn't cause an increase in intensity while exercising, the presence of the video game component kept users engaged for a longer duration which equated to more calories being burned (Geelan et al. 2016).

2.1.1.1.7 Food Force

Some serious games are created with the purpose of raising awareness of specific problems within the world. *Food Force* is a game published by the United Nations World Food Programme (WFP) that aims to spread knowledge and understanding of the procedures associated with humanitarian aid (UNWF 2005). The game is played by managing and balancing the needs of a hungry citizenry with the budget of resources (food) available. Gameplay is divided across several missions that involve locating hungry citizens, producing balanced dietary packs (as seen in Figure 11), air dropping food supplies, coordinating world-wide supplies, leading food convoys, and developing a village over a 10 year span.

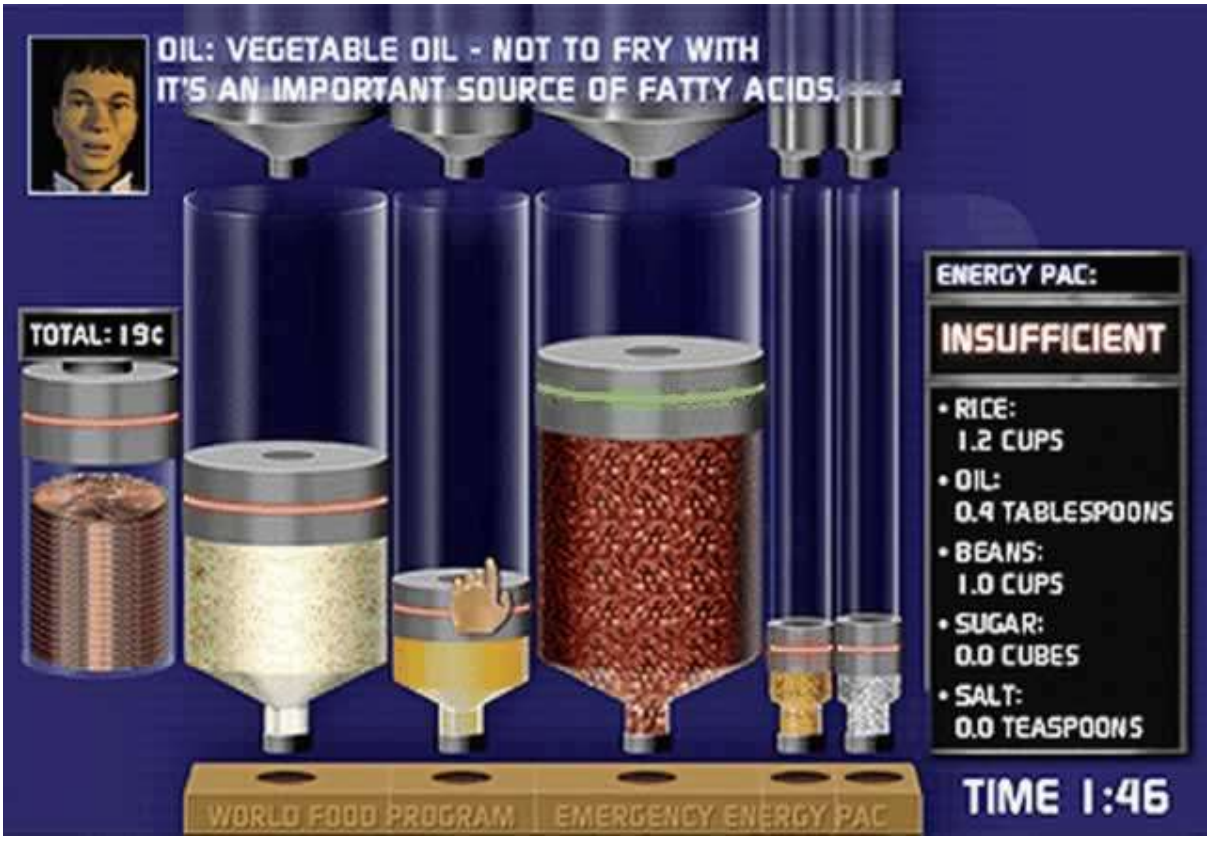

Figure 11 – Food Force energy pac creation

A case study carried out with primary education students found that the game was successful at its goals of helping students learn about emergencies and dispatching humanitarian aid as well as raising awareness of global hunger (Provelengios & Fesakis 2011). Although *Food Force* provided no significant difference in terms of knowledge acquisition as compared to modern pedagogical interventions, the material being presented in a game format significantly affected the engagement of the students during the learning process (Provelengios & Fesakis 2011).

2.1.1.2 Summary of Existing Serious Games

This section has discussed several prominent examples of serious games and how they have positively impacted the space they are based in with highly varied usage from army operations simulation to protein structure manipulation. While the list that has been presented is intended to be a broad representation of the areas in which Serious Games are used, it cannot be understated that the actual breadth of research in the area is much larger than has been showcased here with efforts being put into numerous other projects (Van Lankveld, Schreurs, et al. 2011; Stege et al. 2012; Brewer et al. 2011).

The games themselves also showcase a wide variety of mechanics to from first-person controls to menu-based management games. While serious games are powerful tools that can be used to get players to act in a desired way, they also hold the potential to provide invaluable information via observation of player behaviour in controlled scenarios under the control of designers and researchers.

Video games operate fundamentally by presenting the player with situations of interest and asking the player to act on it. This simple relationship has the potential to reveal much about the decision making process in general. The discussion then turns towards the question of what this data could mean and perhaps the most meaningful way to situate this data is in the field of personality which is elaborated in the following section.

2.2 Literature Review: Personality

One area of research where the concept of *serious games* is of particular interest is the space of personality psychology. Personality refers to individual differences in characteristic patterns of thinking, feeling and behaving (Kazdin 2000). The study of personality focuses on two broad areas of understanding: individual differences in particular personality characteristics, such as sociability or irritability; and how the component parts of a person combine to form the individual ("American Psychological Association" 2015).

As this chapter intends to show, the area of personality research has already utilized video games in various capacities to notable success and that developing a serious game situated in the area of personality research is a logical step to take. Before creation of such a game, in order to gain an understanding of the field of personality, its uses, measures, and limitations a literature review was undertaken.

Six top-ranking journals in the area of general personality research were identified, investigated and analyzed: Journal of Personality, Journal of Personality & Social Psychology, Journal of Research in Personality, Social Psychological and Personality Science, Journal of Personality Assessment, and Social Behaviour and Personality. This analysis, yielded 147 papers that discussed the topic the advancement or uses of identification of personality. The findings are summarized in the following sections.

2.2.1 Popular Models of Personality

This section will highlight the models of personality that appeared more frequently in both a commercial (Pittenger 2005) and academic context (as shown by the literature review).

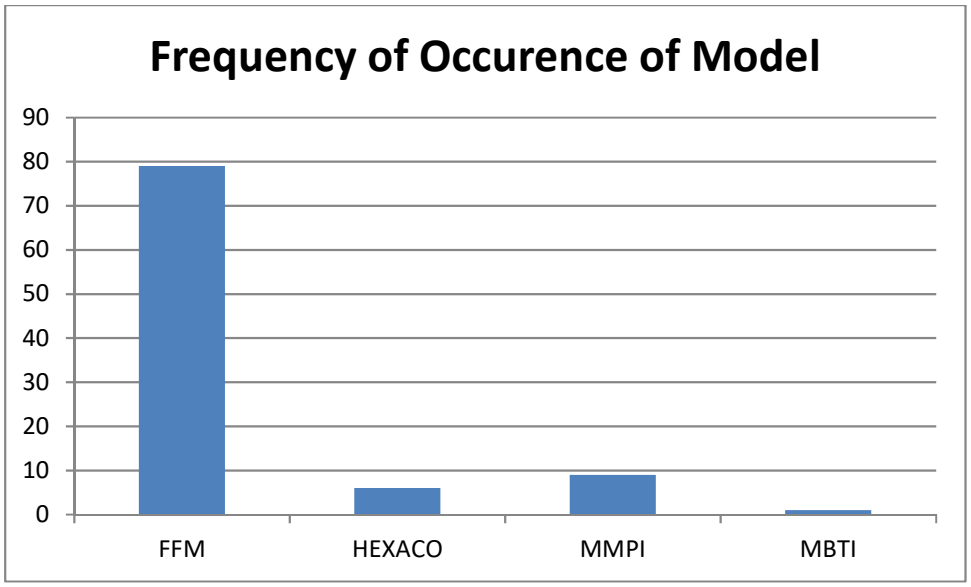

Figure 12 – Frequency of Occurence of Personality Model in Literature Review

Figure 12 highlights the frequency at which each model of personality was mentioned in the literature review. Of the 147 papers reviewed, 79 utilized the Five Factor Model (FFM) (see Sections 2.2.1.3.3 and 3.2) and 6 used the HEXACO model (see Section 2.2.1.3.4) which are both factored models of personality (see Sections 2.2.1.3 and 3.1.3). This review was focused primarily on the general streams of psychology and found that a majority of the papers published in this area used the FFM as opposed to the other models discussed in this section (as seen in Figure 12). A likely reason for this is that the other models tend to have a particular niche, and are not as applicable in the general research area. For example, the Minnesota Multiphasic Personality Inventory (MMPI) (see Sections 2.2.1.2 and 3.1.2) which was developed for, and is used in, more clinical settings such as psychiatry.

The following sections will briefly cover the concept of each of the models of personality identified in Figure 12.

2.2.1.1 Myers Briggs Type Inventory (MBTI)

The Myers Briggs Type Inventory (MBTI) was developed by Katherine Briggs and her daughter Isabel (later becoming Isabel Briggs Myers) who had a long standing fascination

with human behaviour. They formulated some basic theories on individual differences based on their own observations and were surprised when they happened across Jung's work and found similarities between the two. They then continued their work on human behaviour anchored in Jung's theories (Isachsen & Berens 1998).

In the early 1940s, they developed a simple indicator that measures psychic functions and attitudes, allowing people to gain better insights into their own psyche. The Myers-Briggs Type Indicator (MBTI) emerged as a new methodology of comparing and comprehending, differences in human behaviour which claims that no amount of intent or will is able to change the attributes described by the model (Isachsen & Berens 1998).

The MBTI describes 4 scales that combine to give each person one of 16 possible types. These scales are given in Table 1.

Sensing (S)	Intuition (N)
Thinking (T)	Feeling (F)
Judgement (J)	Perception (P)
Extraversion (E)	Introversion (I)

Table 1 – The 4 scales of the MBTI

Those paired scales are then combined to describe a personality profile within the MBTI that describes a person's tendencies in terms of their behaviours and thought processes as shown in Table 2.

ISTJ	ISFJ	INFJ	INTJ
ISTP	ISFP	INFP	INTP
ESTP	ESFP	ENFP	ENTP
ESTJ	ESFJ	ENFJ	ENTJ

Table 2 – The 16 personality profiles of the MBTI

Individuals are grouped into one of the 16 profiles shown in Table 2, which is typically done by completing a questionnaire (see Section 2.2.2.2). The MBTI has found some use in research, for example: accordance with pedagogy and different learning styles (Muller & Pennington 2014). However, the MBTI is primarily used in corporate settings for purposes such as recruitment and team building, offering guidelines and advice on how to approach

interactions with individuals of specific personality types whether it be within a team or between subordinates and superiors. An estimated two million Americans a year take the MBTI with 89 companies out of the US Fortune 100 making use of it (Dattner 2008; Stromberg 2014).

The MBTI's success in a corporate setting is supported by a number of professional organizations that provide a variety of services. The Center for Applications of Psychological Type offers services such as training for administration and interpretation of the test, assistance with scoring, and maintaining a database of MBTI profiles (Pittenger 2005). The Association of Psychological Type (APT) represents the interests of professionals who use the MBTI and also provide workshops that train non-psychologists to purchase and administer the MBTI in nonclinical settings (Pittenger 2005). The MBTI also faces a lot of criticism in research which is discussed in more detail in Section 3.1.1.2.

2.2.1.2 Minnesota Multiphasic Personality Inventory (MMPI)

The Minnesota Multiphasic Personality Inventory (MMPI) is among the most widely used standardized psychometric test of adult personality and psychopathology acting as the focus for a large amount of academic research (Camara et al. 2000; Butcher & Williams 2009), however only appearing in nine out of the 147 articles reviewed (see Section 2.2.1). This is likely due to the model being originally developed for use in medical and psychiatric settings (Butcher & Williams 2009) whereas the focus of the literature review was more concerned (and therefore targeted journals) with more general uses of personality. The specifics of the model are elaborated upon further in Section 3.1.2.

The MMPI was developed by psychologist Starke Hathaway and psychiatrist J.C. McKinley who carefully selected items that covered the symptoms of patients in psychiatric and medical clinics and then insisted on rigorous research during the scale's early life beginning around 1941 (Butcher & Williams 2009). The scale experienced a number of scale additions over the proceeding decades, but the biggest revision since its conception would be in 1989 when the MMPI-2 scale was released which restructured the inventory to trim out extraneous scales (Butcher et al. 1190; Butcher & Williams 2009). Since the MMPI-2, most additions to the scale have been validity scales that are not designed to change core scale assessment (Butcher & Williams 2009).

While the MMPI is the focus for a large amount of academic research (described later in Section 3.1.2.1), its biggest use is in forensic settings (Pope et al. 2006) where the MMPI remains among the measures most commonly applied by forensic psychologists (Archer et al. 2006). A large reason for using the MMPI in the criminal justice system is the broad range of validity scales inherent in the test which can be used to gauge a test-taker's approach to a high-stake evaluation (Ben-Porath 2013). The MMPI is designed to account for factors such as lack of motivation, low reading and language comprehension skills, limited intellectual resources, and cognitive impairment (Ben-Porath 2013). Further, the MMPI is also designed to identify over-reporting of psychological problems as many suspects who take the test may seek to avoid criminal charges or obtain valuable psychotropic medication in the prison system (Ben-Porath 2013).

The MMPI's focus began very clinically and this has been reflected in the uses that the field has found for it and the great effort that was placed on its validity scales also means that sensitive areas such as forensics is able to rely on a personality measure that is hard to fool.

2.2.1.3 Factored Models of Personality

In 1961 Tupes and Christal found five recurrent factors while analysing personality in eight different studies that spanned different populations (Tupes & Christal 1961). This surprised personality theorists of the time as they were far from reaching consensus with no single theory reaching dominance (McCrae & John 1992). Despite its surprising results, not much attention was paid to this finding at the time. In the 1980s, researchers from different schools of thought concluded that these factors were fundamental dimensions of personality and that these five factors were recurring across self-ratings, natural languages, and theoretically based questionnaires, and also across age groups, gender, and language (John et al. 1984).The five factors were shown to have convergent and discriminant validity across instruments and observers and that they endured across decades of adulthood (McCrae & Costa 2012). The established models are the result of differing paths of research over a number of decades that lead to an agreement over a set of common factors. While many different lines of inquiry were undertaken, there are two broad categories for the approaches that researchers use when working with personality: the lexical approach which attempts to identify similarities in

language between people and the more theoretical approach of designing questionnaires based on personality theories.

2.2.1.3.1 The Lexical Approach

The lexical approach to personality model development is based on the hypothesis that all important individual differences will have been noted by speakers of a natural language at some point in the evolution of the language and encoded in trait terms. Thus, by decoding these terms, we can discover the basic dimensions of personality(McCrae & John 1992). As testament to this Allport and Odbert noted some 4,500 trait terms in the English language alone which demonstrates the natural focus by which language evolved to convey aspects of personality (Allport & Odbert 1936).

Although this approach began finding universal trait factors as early as 1936, it wasn't until the 1980s that the Five Factor Model (FFM) (see 2.2.1.3.3) was firmly placed in the mainstream of personality psychology where it has remained until today (McCrae & John 1992; Digman & Takemoto-Chock 1981; Goldberg 1981).Interestingly, similar factor structures emerged when English scales were translated into German, Japanese, or Chinese (Borkenau & Ostendorf 1990; Bond et al. 1975; Yang & Bond 1990) which further alludes to a universal personality model across cultures.

2.2.1.3.2 The Questionnaire Approach

The questionnaire approach to the study of personality works from the theory of personality into a practical application where it is tested via administration to a live population, designed for specific practical applications and measurements (Goldberg 1971). While different theories produce different scales, a number of studies across several decades found distinct universal factors between them which are generally accepted today as universal personality factors (Eysenck & Eysenck 1975; Tellegen & Atkinson 1974; Costa & McCrae 1976; McCrae & John 1992).

2.2.1.3.3 The Five Factor Model

Currently, the most prominent and widely used model that is used for the study of personality within academia is the Five Factor Model (FFM) as evidenced by a review of the field which found that 79 of 147 studies in the past 5 years used the FFM (as discussed in Sections 2.2.1 and 3.2). The 5 factors of personality as captured by the Five-Factor Model are Extraversion,

Agreeableness, Conscientiousness, Neuroticism, and Openness. Each of the five personality factors represents a range between two extremes where most people lie somewhere in between the two polar ends of each dimension (Cherry 2015b). Due to its roots in the lexical approach to personality psychology, each of these dimensions are strongly associated with the trait adjectives that are frequently used to describe the trait (John & S Srivastava 1999).

Each factor in the FFM is itself made up of facets and represents a spectrum of human behaviour and individuals may score anywhere along that spectrum with high and low scores having different connotations for their approach to everyday life and problems (Cherry 2015b).

2.2.1.3.4 Six Factor Model (HEXACO)

The six factor HEXACO model of personality appeared in 6 of 147 articles in a review of the general stream of personality psychology research over the past few years as shown in Section 2.2.1. Although not as popular as the five factor model, this HEXACO model has still been the focus of some research (Nel et al. 2012; Klimstra et al. 2014). The six factors of the model as described by the HEXACO abbreviation: Honesty-Humility, Emotionality, eXtraversion, Agreeableness, Conscientiousness, and Openness to experience is very similar to the five factor model (FFM) (see Section 2.2.1.3.3). The HEXACO model the Extraversion, Agreeableness, Conscientiousness, and Openness scales in similar terms, but foregoes the Neuroticism scale in favour of the additional Honesty-Humility and Emotionality scales.

2.2.1.3.5 Usage of the Factored Models of Personality

The factored model of personality offers insight into a large spectrum of the human psyche. While the debate of whether or not this insight is complete (or complete enough) goes on (John & S Srivastava 1999; Gnambs 2014; Woo et al. 2014), the models themselves have been used in a wide array of research that bridges many varied fields with personality psychology in order to add new perspectives to the collective body of knowledge (John & S Srivastava 1999).

For example, the lexical origin of the factored personality model makes it an ideal tool for the purpose of research between different cultures and ethnicities. By searching for universal aspects of personality across languages, we are also able to develop better understanding of

universal traits, values, and behaviours shared by geographically distinct (McCrae et al. 2010; Tackett et al. 2012; Gurven & Rueden 2013).

Since personality itself refers to a relatively stable pattern of behaviour, affect, and thinking (Saucier et al. 2013), it stands to reason that there also exists a fair amount of behavioural research where the 'how' and 'why' of human decisions are picked apart from the perspective of personality (Hilbig et al. 2014; Quilty et al. 2014; Dobewall et al. 2014). This line of research aims to create frameworks that support self-reflection to understand our own decisions and also give us insight into other's decision making to foster understanding between individuals. In furthering that goal, the factored personality models have also found applications in categorizing and understanding nuances in relationships (Furler et al. 2014; Nezlek et al. 2011).

Bridging the gap with health, the factored personality models have had and continue to have a massive impact with its promise of great predictive power (Widiger & Presnall 2013). Specifically, some work has been put into tackling problems associated with depression and anxiety (Lewis et al. 2014; Chow & Roberts 2014), cardiovascular risk (Gleason et al. 2014), and metabolism (Human et al. 2013; Israel et al. 2014) among others. This line of research holds the potential for early detection of complicated health problems that can be mediated by lifestyle changes and save money and lives.

Unsurprisingly, the factored personality models have also been the basis for many contributions to our knowledge of personality disorders (Gleason et al. 2014; Widiger & Costa 2012). Importantly, the FFM has been used heavily in concord with the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5) which is the American Psychiatric Association's (APA) classification and diagnostic tool which is a critical tool in the diagnosis and subsequent treatment of mental problems (Trull 2012; Widiger & Presnall 2013; Mullins-Sweatt & Lengel 2012).

In addition, the factored personality models have been used in several longitudinal studies that aimed to explore the link between early life development and intellect scores (Chamorro-Premuzic et al. 2004; DeYoung et al. 2014; Abe 2005). Those studies showed promise for the use of early personality testing to predict varied factors such as self-regulation and academic

performance which can lead to early interventions for poor developmental habits (Chamorro-Premuzic et al. 2004; DeYoung et al. 2014; Abe 2005).

All of the examples above serve to illustrate the extremely varied and useful applications of this line of research into understanding the fine nuances of human behaviour.

2.2.2 Personality Elicitation Tools

The previous section discussed the various models of personality that were discovered during a review of personality psychology as described in Section 2.2.1. This section intends to discuss the tools used in the personality information elicitation process. The review of the field found that questionnaires are the primary form of data collection in the area of personality research as summarized in Figure 13 below.

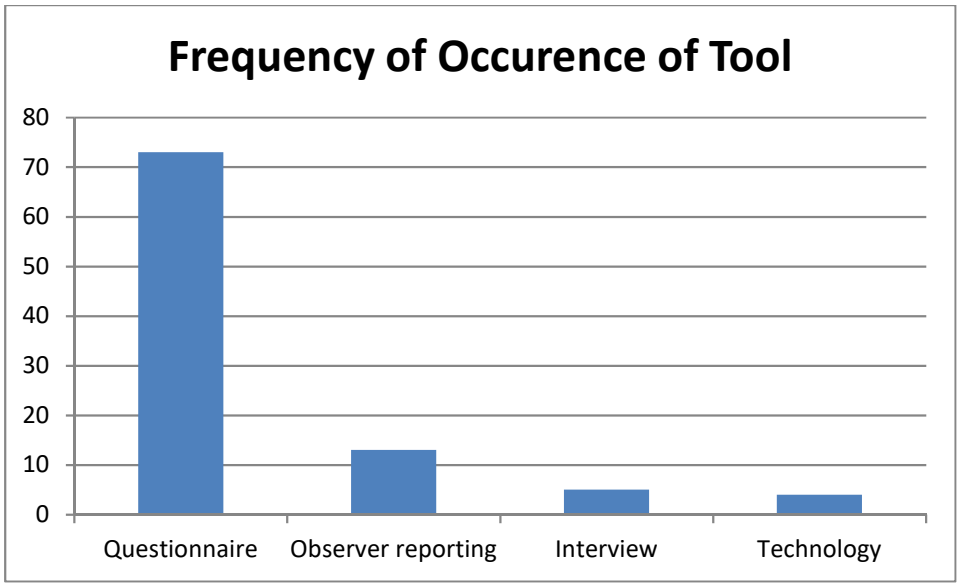

Figure 13 – Frequency of Data Elicitation Tools in a Review of the Field

In a review of 147 papers over the past 5 years, 73 used questionnaires (see Section 2.2.2.2), 13 used observer reports (see Section 2.2.2.4), 5 used interviews (see Section 2.2.2.3), and 4 used various technology-based tools to sample personality data (see Section 2.2.2.6). Questionnaires were used more than any other tool for the purposes of personality elicitation. The following sections will first discuss problem of Social Desirability Bias (SDB), which is an inherent problem with gathering self-reported personality data, and then examine the properties of each of the tools shown in Figure 13 as well as some of the problems inherent in those methods.

2.2.2.1 Social Desirability Bias

Here it is important to note that all explicit tests of personality are vulnerable to Social Desirability Bias (SDB) which is the tendency for people to present a favourable image of themselves (Van Lankveld et al. 2009; Grimm 2010; King & Bruner 2000; Mortel 2008). As long as the personality profile being generated relies on information that is volunteered by the target, the results will be influenced by SDB to some degree (even when participants aren't aware of it) and the tool will have to address this issue in some manner (King & Bruner 2000; Grimm 2010). Questionnaires (see Section 2.2.2.2) in particular have to take SDB into account as they rely on self-reports and participants to be as honest as possible (King & Bruner 2000).

The issue of SDB influencing the results of a questionnaire has been scrutinized in great detail by many researchers. Some look at the effects of SDB in specific fields such as personality testing with personnel selection or astronaut selection (Ones et al. 1996; Sandal et al. 2005) while many more have articulated problems with SDB in a more general way of affecting data elicitation, showing that while few studies detect or control for SDB, almost half of those that did found SDB influencing their results (Neeley & Cronly 2004; Grimm 2010; Fisher 1993; Parmač Kovačić et al. 2014; King & Bruner 2000).

As a result, research has been put into the task of detecting, reducing, and eliminating SDB (Fisher 1993; Nederhof 1985) including adding a set of extra questions that measure SDB (Fisher 1993) or to design a survey environment where the effects of SDB are mitigated (Nederhof 1985). An example of the former is the use of structured, projected questioning where the respondent is asked to comment on what they think others do (and thus depersonalize themselves from the question) (Fisher 1993), but that does not work in the context of a personality test since the respondent is required to provide intimate and personal details about themselves (Grimm 2010). The latter ranges from carefully balancing the choices in a questionnaire to be equally socially desirable (which is a subjective matter and difficult to moderate) to controlling the environment in which the questionnaire is answered to pressure the respondent to answer more truthfully, making them think that they cannot lie using a pseudo-lie detector (Nederhof 1985).

Personality inventories commonly use Crowne and Marlowe Social Desirability Scale (or similar questions) to measure the rate of SDB in a respondent (Crowne & Marlowe 1960). The

scale asks participants to rate themselves on a set of behaviours that are perceived by society to be exemplary, but only enacted infrequently (Loo & Thorpe 2000). If a respondent associates themselves with many of these behaviours, they are likely to be exhibiting SDB (Moss 2008). The problem with this method is that the additional questions bloat the total size of the questionnaires and may cause fatigue in respondents which reduces the accuracy of the data anyway (Cape 2010). Thus, it becomes a balancing act of keeping the number of questions as small as possible to reduce fatigue but also using enough such that the self-deception error can be minimized.

At this point, it is important to distinguish between implicit and explicit measurements of personality. Implicit measures give an idea of how the individual in question is perceived by those around them as the inferences on their personality are made based on the actions that they are making regardless of the intent behind them while explicit constructs (including any built-in SDB) give us an idea of the individual's own perception of themselves or even their desired self-image (Boldero 2007). Looking at this idea in another way, the implicit measurements would be how the target is perceived by others and the explicit measurements are how the target reports themselves.

2.2.2.2 Questionnaires

The traditional questionnaires personality research relies upon tend to either be Likert scale or true/false items and range from anywhere between 44 to 536 items (John & S. Srivastava 1999; Butcher & Williams 2009). In a review of 147 papers over the past 5 years, 73 used questionnaires as shown in Section 2.2.2. As questionnaires ask respondents to report on their own thoughts, they can be considered explicit measures of personality.

Personality researchers have explored alternative data collection techniques with varying amounts of success (as will be explored in the coming sections) in the forms of interviews and observer reports (Mund & Neyer 2014; Letzring & Human 2014; Soto & John 2014). These techniques address the issue of SDB by taking the onus of judging personality away from the target in question and instead relies on third parties (Mund & Neyer 2014; Letzring & Human 2014; Soto & John 2014).

2.2.2.3 Interviews

The concept of personality within a person is somewhat subjective in that it attempts to formalize the behaviours and thought processes of an individual (Kazdin 2000). The questionnaire approach asks that the target individual is able to communicate those factors by answering targeted questions (John & S Srivastava 1999). In contrast, several studies have shown that a second party is also able to form a personality profile based on an interaction with the target in question (Chen et al. 2014; Gleason et al. 2014). As personality profiles developed via interviews are based on the impressions of the interviewer, they can be considered implicit measures of personality.

Just as questionnaires face the inherent weakness of SDB, the concept of using interviews face an inherent weakness of subjectivity (Chen et al. 2014). The interviewer in this situation is making value judgements on the actions and reactions of the target, limiting the generalization between interviewers and potentially opening up doors to other biases that are let in by human judgement calls (Chen et al. 2014). Further, the interview process itself is a time consuming process which requires more resources as compared to traditional questionnaires (Chen et al. 2014).

2.2.2.4 Observer Reports

Another alternative to the traditional questionnaire method is the use of observer reporting which were used in 13 of 147 studies reviewed over a five-year period. This approach suppresses the effects of SDB by using reports obtained from the friends or family of the target (McCrae et al. 2010; McCrae & Terracciano 2005). The idea is that the problems inherent with personality judgement are mitigated by the fact that it is easier to be objective about someone else. The notion has further been extended to complete strangers making assessments of the target with positive results (Connelly & Hülsheger 2012; Human et al. 2014; Israel et al. 2014). As these personality profiles rely on the impressions of a third party, observer reports can be considered implicit measures of personality.

Interestingly, there have been findings that observer reports are consistent enough to be able to predict certain health outcomes of young adults as they enter midlife(Israel et al. 2014). In this case, the observation report was completed by the hospital staff after interaction with the targets who were part of a long-term longitudinal study examining the effects of personality

on long-term health(Israel et al. 2014). Within this context, a very brief 'exposure' period was enough to obtain a consistent and meaningful result (Israel et al. 2014).

The observer reports tend to simply be standard personality questionnaires that have been retooled to describe the target instead of the test-taker (Connelly & Hülsheger 2012; Israel et al. 2014). These ratings take a middle road between standard questionnaires and getting professional psychologists to observe the target, with some research that indicates it may be a better measure for certain traits (Vazire 2010).

In general, while the observational methods are considered to be more reliable and objective than self-reports, they suffer from a higher cost and requiring more effort (Arney 2004; Van Lankveld et al. 2009).

2.2.2.5 Media Preferences

As a number of personality models claim to capture information about human behaviour in a fairly general and universal way, it stands to reason that some of that will be reflected within specific choices that are made every day. For example, several studies have also shown that media consumption preferences offer enough information to be able to draw inferences about one's personality.

2.2.2.5.1 Traditional Media Preferences

The studies in question have found correlates between preferences for film, television, music, and books with personality traits (Cantador et al. 2013; Rentfrow, Goldberg & Levitin 2011; Rentfrow, Goldberg & Zilca 2011). These works in general clustered the responses provided to them into distinct groups depending on the medium in question. The groups that arose from that analysis were then correlated to the personality traits of the FFM (Cantador et al. 2013; Rentfrow, Goldberg & Levitin 2011; Rentfrow, Goldberg & Zilca 2011). The results of these studies show that the nuances of choice are able to be captured through the consumption of media.

2.2.2.5.2 Video Game Preferences

Another specific instance of media being used in conjunction with personality testing is a study that sought to examine motivation, play styles, and preference of video game players (DeGraft-Johnson et al. 2013). This was done by correlating five factor personality scores with

how much someone liked a game and how difficult they found it across four genres: fighting, racing, dancing, and first person shooter (FPS) games (DeGraft-Johnson et al. 2013) on the PlayStation 3 and Xbox 360 consoles (Sony 2015; Microsoft 2015).

Genre	Platform	Games
Fighting	PlayStation 3	Tekken Hybrid
		Soul Calibur IV
Racing	PlayStation 3	Grid
		Burnout Paradise
Dancing	Xbox 360	Just Dance 3
		Dance Central 2
First Person Shooter	Xbox 360	Halo 3
		Call of Duty: Modern Warfare 2

Table 3 – List of Games Used in Personality-Game Preference Correlate (DeGraft-Johnson et al. 2013)

Despite the small range of genres and only using eight games as shown in Table 3, the study reported significant findings for several correlates from its 62 valid participants (DeGraft-Johnson et al. 2013):

- Moderate negative relationship between conscientiousness and ease of playing an FPS (p. 4)
- Positive Pearson correlation between extraversion and ease of dancing games (p. 5)
- Moderate positive relationship between extraversion and liking dancing games (p. 5)
- Positive relationship between agreeableness and liking dancing games (p. 5)

While no more correlations were found due to the scope and limitations, the findings at least suggest that there are general assumptions that can be made about whether or not someone would like a game and how well they make take to it depending on their personality.

2.2.2.6 Technology Based Sampling

The area of personality research has also begun using more technologically grounded techniques to data collection as a means of overcoming SDB, such as examining the usage of

ubiquitous social media like Facebook or microblogs to elicit personality information from users (Bai et al. 2013; Park et al. 2014). These techniques bypass the problems of SDB by not asking the respondent to answer questions, but rather by observing patterns that arise from their usage of various technologies. The following sections will discuss some of the technology sources that have been explored for links with personality data elicitation.

2.2.2.6.1 Smartphones

A study published in 2013 reported success when passing smartphone usage data through an algorithm to produce a personality profile (Chittaranjan et al. 2013). The study used data from 117 Nokia N95 smartphones collected over a continuous period of 17 months in Switzerland (Chittaranjan et al. 2013). The researchers developed a machine learning method to detect personality traits of a user based on their smartphone usage (Chittaranjan et al. 2013). The study laid the basis for the concept of using predictive technology to perform automatic analysis on user information in order to derive a personality profile (Chittaranjan et al. 2013).

2.2.2.6.2 Facebook

Further building on the idea of using information from a common everyday tool for personality data collection, researchers had claimed that not only does one's Facebook profile reflect their reflected (implicit) personality as opposed to their self-idealized (explicit) personality.

A prominent study used a popular Facebook application and managed to develop an algorithm that processed the Facebook behaviour of a target and produced a personality profile with a high rate of success and similarity to the results of the traditional questionnaire (Park et al. 2014). A technique such as this would directly observe an individual's behaviour and theoretically overcome any SDB so long as the target was unaware they were being observed.

2.2.2.6.3 Video Games

The study of personality has also crossed over with the field of video games (Shen et al. 2012; Canossa et al. 2013; Van Lankveld, Spronck, et al. 2011; Van Lankveld, Schreurs, et al. 2011; Spronck et al. 2012). The popular massively multiplayer role-play game (MMORPG) World of Warcraft (WoW) (Inc. 2015) was used as a lens of capturing player behaviour and information which was then used to infer personality information successfully (Shen et al.

2012). This assessment was focused on behavioural traces, textual data, and social networking information present within the game environment (Shen et al. 2012). Those data points and any kind of information that describes an action taken by the user is known as a metric and is similar to the concept of using Facebook as a platform to measure player personality as an MMORPG is a digital environment that generally requires a lot of time commitment where many social bonds are created and nurtured. Within a game, any and all actions undertaken by the player can conceivably be measured as a metric for future analysis and is a frequently used source of data within the games industry to improve the quality of their games (Mason 2014).

In another instance, the game Minecraft was used as the focal point of a study that sought to better understand player psychology, behaviour, and motivation (Canossa et al. 2013). The game itself ostensibly allows players to create anything from sets of cubes that look and behave differently, where some from around the world have shown off creative and impressive feats such as creating a fully functional 16-bit computer and a programmable piano out of simple switch circuitry (Ohmgane3sha 2011; FVDisco 2011). The study took advantage of that freedom of play style and range of possible metrics to monitor interesting and useful information about one's psyche (Canossa et al. 2013). In a similar vein, the game Fallout 3 (Bethesda 2009) was used for its introductory sequence in order to correlate observations on player behaviour with their personality profiles (Spronck et al. 2012).

The games WoW, Minecraft, and Fallout 3 already offer a large range of metrics to utilize that would inform personality assessment. In those cases, and even the studies that use Facebook and smartphones, a researcher may only harvest incidental data from participants as measures towards personality. This is a very powerful approach as it draws many parallels with the use of observer reporting and overcomes the issues of SDB, as well as being easier to organize as there is no coordination with extra people to gather data from (Mortel 2008). However these methods come with their own set of weaknesses (Canossa et al. 2013; Shen et al. 2012; Spronck et al. 2012). Firstly, it relies on access to the measures themselves which may be obfuscated and difficult to get to or even simply not available to the public or academic members. Secondly, there is no control over the behaviour of the underlying system that the measurements take place on in the event that researchers seek to make changes to the stimuli

being presented (Canossa et al. 2013; Shen et al. 2012; Spronck et al. 2012). Thirdly, these media only offers accurate readings on those who use the system on a daily basis (Chittaranjan et al. 2013; Park et al. 2014). Should research require information from those who do not use the platform in question, then starting at the need of the research would poison the validity of the results obtained since it is now an action initiated by the researcher (Shen et al. 2012; Sheng-Yi et al. 2012).

In contrast to the methods described above, a study at the University of Tasmania in 1999 created a game environment for the purpose of agoraphobia treatment that found behavioural correlations between certain in-game actions and certain personality traits (Kirkby et al. 1999). This game was not a commercial entertainment game, but rather one developed with the serious purpose of addressing agoraphobia and yet was able to showcase enough of a range of possible actions and metrics to claim significant certainty of personality trait correlation (Kirkby et al. 1999).

Taking that idea a step further, two studies created a virtual environment for the purpose of personality assessment with a video game scenario that was designed using the pre-existing Neverwinter Nights (BioWare 2002) which is a modifiable role-play game (RPG) (Van Lankveld, Spronck, et al. 2011; Van Lankveld et al. 2009). In both studies, researchers custom-built the story, world, characters, and quests for the player to interact with and found that the player choices recorded in the video game matched specific personality traits and could be used to create a complete personality profile (Van Lankveld, Spronck, et al. 2011; Van Lankveld et al. 2009). The scenarios used in those video games ask the player to make decisions in a menu-based selection, choosing their course of action from a list of pre-written answers (Van Lankveld, Spronck, et al. 2011; Van Lankveld et al. 2009). As compared to the use of existing games like WoW, Minecraft, or Fallout 3, or networks like Facebook, this gives full control over the measure of personality and turns the video game into a much more powerful tool for personality elicitation (Van Lankveld, Spronck, et al. 2011; Van Lankveld et al. 2009).

These examples show that a game is able to confront a player with specific situations where they make decisions and possibly offer direct insight into their thought processes which could be used by researchers to create personality profiles (Shen et al. 2012; Van Lankveld, Spronck,

et al. 2011; Canossa et al. 2013). Further, in those contexts the video game is able to directly record and measure the actions and behaviours (Shen et al. 2012; Van Lankveld, Spronck, et al. 2011; Canossa et al. 2013). This is comparable to the questionnaires which ask participants what they think of or what they would do in certain scenarios but instead providing a context where they can act upon the situation in question (Van Lankveld, Spronck, et al. 2011; Van Lankveld et al. 2009).

2.2.2.6.4 Summary of Technology Based Sampling

The examples discussed above suggest that the technology shows potential for the creation of a method of personality assessment that is virtually instantaneous (Van Lankveld, Spronck, et al. 2011) with a failing that it requires consistent use by the user (Park et al. 2014; Chittaranjan et al. 2013). Since the source of data for personality evaluations come from the usage of a specific tool, if for example the Facebook application was interfacing with a barely used Facebook profile, there would simply not be enough information to form an accurate reading (Park et al. 2014; Chittaranjan et al. 2013). However, games have also shown potential to provide an even greater amount of information and insight into the psyche (Shen et al. 2012; Van Lankveld, Spronck, et al. 2011).

2.2.2.7 Summary of Personality Tools

There are a number of personality data collection methods such as the traditional questionnaire which dominates the field (see Section 2.2.2.2), interviews which rely on subjective judgement of a second party (see Section 2.2.2.3), or using separate observers to rate the target (see Section 2.2.2.4). A recent trend has shown the use of newer technologies as a means of data elicitation which shows potential to not only match the results from traditional methods (Chittaranjan et al. 2013; Park et al. 2014), but also provide novel and interesting insights into personality (see Section 2.2.2.6).

A rich source of this data comes from social networks such as Facebook or even the social structure found in online communities (Park et al. 2014). This is limited by the fact that these measures require consistent use of the system by the user as the algorithm draws context for personality evaluations from the established behaviours on the platform (Park et al. 2014).

Building upon the idea of extracting data from a user as they interact with a system, a video game holds the potential to provide an even greater amount of information and insight by placing targets into a more literal scenario as compared to traditional questionnaires (Van Lankveld, Spronck, et al. 2011; Van Lankveld et al. 2009; Van Lankveld, Schreurs, et al. 2011) and measure behaviours and reactions directly instead of through the abstraction of questions in a survey (Shen et al. 2012; Canossa et al. 2013). This observational behaviour recording operates on the same fundamental concepts of observer reports (see Section 2.2.2.4) and potentially holds the ability to circumvent problems with SDB without the drawbacks of logistically handling multiple individuals during a research period (Shen et al. 2012; Van Lankveld, Spronck, et al. 2011; Canossa et al. 2013).

At its core, video games are simply set scenarios that players are reacting to (Moore 2011). Thus, it follows that if those scenarios result in meaningful data for the assessment of personality, video games could provide a unique lens at the phenomena of personality (Shen et al. 2012). The intersect between these two areas is an exciting opportunity space.

2.3 Player Motivation

It should be noted that personality types alone do not describe the full range of motivation behind player actions. As interest in video games has increased a host of research has been placed into dissecting human decision making processes within video games (Canossa et al. 2015; Tekofsky et al. 2015; Tekofsky, Spronck, Plaat, Herik, et al. 2013). There have been strong arguments made that individual personality relates weakly to player behaviour and that other factors such as age provide a stronger correlation (Tekofsky et al. 2015; Tekofsky, Spronck, Plaat, Van Den Herik, et al. 2013). However, based on the evidence provided in Section 2.2.2.6.3, it should be clear that given the right circumstance there are interesting and important correlations to be examined.

Further, research that was published after the initial literature review phase and was only discovered after too much work had been put into the game's design (Section 3.4) to be revised highlighted correlations between traits and behaviours were specific to game areas that carry different situational affordances (Canossa et al. 2015). The implications of this research to the design process is discussed in Section 6.3.

2.4 Opportunity for Research

Serious games have been used in conjunction with many fields in order to achieve a purpose other than entertainment (see Section 2.1.1.1). Their power lies in the motivational and engaging nature of video games (see Section 2.1.1). An opportunity exists to explore the idea of using serious games in a strictly observational way as opposed to one whose purpose is educational (see Section 2.1.1.1.2 or 2.1.1.1.7) or to foster behaviour change (see Section 2.1.1.1.6 or 2.1.1.1.5) and in doing so gain a better understanding of player behaviour (see Section 2.1.1.2).

This pairs logically with the area of personality research where the primary tool for data collection is the questionnaire which is hindered by SDB (see Section 2.2.2.2). Third party information has been found to overcome the problems posed by SDB (see Section 2.2.2.4), with research even extending to utilizing computers to process behavioural information to create personality profiles (see Section 2.2.2.6). Many of the examples discussed utilized existing technologies such as the social media platform Facebook (see Section 2.2.2.6.2) or video games (see Section 2.2.2.6.3), but some also utilized custom designed video game scenarios in order to analyze player behaviour (Van Lankveld, Spronck, et al. 2011; Van Lankveld et al. 2009).

While the (Van Lankveld, Spronck, et al. 2011) and (Van Lankveld et al. 2009) studies utilized custom scenario building, both cases still relied on an existing game engine: Neverwinter Nights (BioWare 2002). Therefore, there is an opportunity here to create a serious game where researchers have full control over the scenarios in the game in order to place players in specific situations and monitor their behaviours. The scenarios in such a tool would be designed from the ground up, incorporating personality theory from the very beginning instead of adjusting an existing tool to the same goal. The hope for the designed video game tool is to avoid explicitly asking players questionnaire-like items but rather to observe their behaviour in the game and look for correlations between gameplay choices and personality. The goal is to recreate instances like those mentioned in Section 2.2.2.6.3 where gameplay trends indicate a certain personality grouping while still offering the designers of the game full creative control and to then learn about the unique insights of creating and designing a video game using an existing theoretical framework as the inspiration.

To that end, a research question is formulated as follows:

What are the opportunities and challenges in creating a serious game to support the identification of player personality characteristics?

This research question can be further broken down into the following sub-research questions:

1. What are the opportunities and challenges in designing a serious game to support the identification of player personality characteristics?
2. What are the opportunities and challenges in implementing a serious game to support the identification of player personality characteristics?

These questions are expanded upon in the following chapter.

3. Methodology

The previous sections concluded with the formalization of a research question that attempts to solve the problem of creating a custom built serious game for the purposes of personality elicitation. This will allow researchers full control over the scenarios that players encounter in order to facilitate a better data elicitation process.

The research question is as follows:

> What are the opportunities and challenges in creating a serious game to support the identification of player personality characteristics?

This research question can be further broken down into the following sub-research questions:

1. What are the opportunities and challenges in *designing* a serious game to support the identification of player personality characteristics?
2. What are the opportunities and challenges in *implementing* a serious game to support the identification of player personality characteristics?

The questions differentiate between the design and the implementation of a video game personality elicitation tool. Sub-research question 1 (SQ1) is concerned with designing the tool in such a way that it is supported by established theory. In order to do so, a model of personality that best suits the project's needs has to be identified. This is achieved through a review of the existing models of personality (see Section 3.1). Next, a process has to be identified that is able to take aspects of the model of personality identified and produce a game design. This is achieved through focused group discussions with subject matter experts from the field of personality as well as video game design (see Section 3.4). As part of that process, good game design principles also have to be adhered to which will be identified through a literature review process.

Sub-research question 2 (SQ2) is concerned with the challenges and opportunities in the process of turning the design derived from SQ1 into a functioning game experience. In order to answer this question, an understanding of how the designed video game matches up to the traditional tools is required. Firstly, this requires a personality profile derived from the traditional tool to be used as a basis for comparison which can be obtained through the

administration of a personality questionnaire (see Section 3.6). Secondly, the data from that questionnaire needs to be compared with the data obtained from the video game tool. Thus, the player's behaviour within the video game will need to be recorded and analyzed (Chapter 4).

Sub-research question	Information Required	Method
1: What are the opportunities and challenges in *designing* a serious game to support the identification of player personality characteristics?	Viable model of personality	Literature review
	A process to convert aspects of the personality model to game design	Focus group discussion with subject matter experts
	Good game design principles	Literature review
2: What are the opportunities and challenges in *implementing* a serious game to support the identification of player personality characteristics?	The personality profile from a traditional measure	Questionnaire
	The results from the game's measures	Gameplay logging

Table 4 – Summary of information and associated method required to answer the research question

Table 4 summarizes the data needed to answer the research and sub-research questions as well as the methods necessary to obtain that data. The goal of this project is to create a video game tool that contains scenarios that have been fully created by researchers in order to measure personality data in players. Ideally, the designed video game tool's result will be comparable to the current dominant method of personality elicitation – questionnaires (as described in Section 2.2.2.2). The following sections describe the work done in order to fulfill the information needs to answer the sub-research questions discussed above.

3.1 Personality Model Selection

This section will discuss the selection of a model of personality for the project in order to begin answering SQ1 which in turn answers the larger research question as shown in Table 4. In order to accomplish this, the popular models of personality described in Section 2.2.1 will be

analyzed for their relevancy to this project. Two criteria will be evaluated for this: the area that the model is used and the validity of the model.

3.1.1 MBTI Relevancy

The MBTI was developed by Katherine Briggs and her daughter Isabel and describes 4 scales that combine to describe each person with one of 16 possible types (see Section 2.2.1.1). The MBTI is a general model of personality (Isachsen & Berens 1998), though has found prominent use in corporate environments (Burnett 2013; Dattner 2008; Stromberg 2014). Despite appearing in only one of the 147 studies examined (see Section 2.2.1), the MBTI was included in this review because of its presence and spending power in the corporate world (Dattner 2008; Burnett 2013; Stromberg 2014) which qualifies it as a model worth considering for widespread use and recognition.

3.1.1.1 Area of Use of the MBTI

As can be seen in Section 2.2.1, the MBTI does not feature heavily in academic literature, although it is the basis of some research in different fields like being used in accordance with pedagogy and different learning styles (Muller & Pennington 2014). The MBTI is primarily used in corporate settings for purposes such as recruitment and team building, offering guidelines and advice on how to approach interactions with individuals of specific personality types whether it be within a team or between subordinates and superiors. An estimate of two million Americans a year take the MBTI with 89 companies out of the US Fortune 100 making use of it (Dattner 2008; Stromberg 2014).

The MBTI's success in a corporate setting is supported by a number of professional organizations providing a variety of services. The Center for Applications of Psychological Type offers services such as training for administration and interpretation of the test, help with scoring, and maintains a database of MBTI profiles (Pittenger 2005). The Association of Psychological Type (APT) represents the interests of professionals who use the MBTI and also provide workshops that train non-psychologists to purchase and administer the MBTI in nonclinical settings (Pittenger 2005).

3.1.1.2 Validity and Rejection of the MBTI

Despite its widespread use in corporate settings, the MBTI is greatly criticized and regarded as a flawed model of personality as indicated by the review of the area (see Section 2.2.1). Several of these validity problems are highlighted in the following sections.

3.1.1.2.1 Flawed Theoretical Foundation

The first prominent means of assessing the validity of a model is its theoretical basis or the claims that it makes and the theory that underlies it. In this context, validity is referring to the degree to which a test measures what it sets out to measure (Pittenger 2005).

3.1.1.2.1.1 Factor Analysis

Factor analysis is a type of statistical procedure that consists of making an analysis of the correlations among the questions in the test (Pittenger 2005). This analysis is undertaken on questionnaires typical to each model of personality and should reveal that there are distinct and separate clusters or factors in the questions being posed which should align with the theoretical factors of the theory of personality in question.

Working with the theory that the MBTI model proposes four dimensions that are unique and stand alone, analysis of the questions should reveal that each question correlates to a single factor and that in total there will be four factors. However, studies have found that this is not the case. For example, a study using 1,291 college aged students found six different factors and a high level of measurement error that lead the authors to conclude that their results were inconsistent with the MBTI theory (Sipps et al. 1985). Other research has also called the Judging-Perceiving and Sensing-Intuition scales into question as they found correlations between one another, thus weakening the MBTI's claim that there are four distinct factors at play (McCrae & Costa 1989).

3.1.1.2.1.2 Predictive Power

Another measure of the theoretical soundness of a model is the predictive power of its measures when compared to its base theory. That is to say that the results of a personality profile should be usable as data to predict certain outcomes relating to the individual or even the population the test is administered on.

For example, the purpose of the MBTI model's existence at all is to predict certain aspects of one's life based on their personality. It has frequently been used in the context of an individual's work and career, claiming to have an amount of insight into the kinds of people who end up in certain professions (Isachsen & Berens 1998). Advocates of the model may paint in broad strokes and claim that certain types dominate certain professions and that the type of an individual is thus telling of what career choices are more comfortable for that person. The problem here is that those claims tend to lack other contextual information that may explain this correlation. For example, nurses tend to have a different distribution of type as compared to managers. While the MBTI type could be the distinguishing factor between the two populations, there are alternative interpretations to that data. For example the fact that nursing has been and remains a profession dominated by women is a much more likely reason for disparity between the groups (Pittenger 2005). That is not to say that the types could not be a factor for profession distribution, but data suggests the proportion of MBTI types within each occupation is equivalent to that of a random sample of the population (Pittenger 2005).

3.1.1.2.2 Tool Reliability

Reliability refers to the consistency of a test between measurements, resulting in similar (or ideally the same) results every time the test is administered (Pittenger 2005). The MBTI claims that a person's type is immutable and doesn't change in their life time (Isachsen & Berens 1998). It would thus be expected that the reliability of the MBTI is extremely high and subsequent retests always yield the same or at least similar results.

The primary method for testing this reliability is to administer the same test to an individual on two occasions with the interval between the test and retest can range between several weeks to more than a year (Pittenger 2005). With the understanding that a person's personality does not change over time, this interval should have little to no change in someone's resulting profile. However, studies have shown that although the proportion of reclassification into the same category can be as high as 90%, the range drops to as low as 50% over a 5-week interval for the MBTI (Boyle 1995; Pittenger 2005).

3.1.1.3 Final Analysis of the MBTI

In summary, the MBTI fails on a number of points in relation to validity: first, whether or not the dimensions described by the theory really exist (investigated via factor analysis); second, whether knowing a person's profile really gives someone predictive power over the target individual's behaviour in different circumstances; and third, whether the results of the testing are consistent over time (Pittenger 2005). In those ways, the MBTI falls short of the mark, is often ignored in academic research fields and will not be used for the purposes of this project.

3.1.2 MMPI Relevancy

The Minnesota Multiphasic Personality Inventory (MMPI) is the most widely used standardized psychometric test of adult personality and psychopathology acting as the focus for a large amount of academic research (Camara et al. 2000; Butcher & Williams 2009). The MMPI was used in nine out of the 147 studies examined in the literature review (see Section 2.2.1). Questionnaires for the MMPI consist of a number of true/false questions (depending on the version of the questionnaire) with each question corresponding to one of 10 clinical scales that each indicate different psychological conditions (Cherry 2015a).

3.1.2.1 Area of Use of the MMPI

The MMPI was originally designed to be used in medical or psychiatric clinics and continues in those fields today (Hunter et al. 2014), however the measure has also received wide use in other fields such as personnel screenings (Butcher et al. 2006) for sensitive jobs like airline pilots, police, or nuclear power plant operators, correctional settings (Sellbom 2014), family custody (Ezzo et al. 2008), and personal injury evaluations(Livingston et al. 2006). The MMPI has experienced much success through being translated into different languages and helping to bridge the gap in terms of understanding different cultures via personality (Butcher & Williams 2009).

Arguably though, its biggest uses is in forensic settings (Pope et al. 2006) where the MMPI remains among the measures most commonly applied by forensic psychologists (Archer et al. 2006). A reason for using the MMPI in the criminal justice system is the broad range of validity scales inherent in the test which can be used to gauge a test-taker's approach to a high-stake evaluation (Ben-Porath 2013). Factors like lack of motivation, low reading and language comprehension skills, limited intellectual resources, and cognitive impairment may

compromise an individual's capacity to respond meaningfully when confronted with a psychological test (Ben-Porath 2013). After all, at the point of forensic psychology, many suspects have a strong motivation to over-report psychological problems such as avoiding criminal charges or obtaining valuable psychotropic medications in the prison system (Ben-Porath 2013).

3.1.2.2 Validity of the MMPI

In addition to the 10 clinical scales, MMPI questionnaires also feature a number of validity scales that have been built into the tool(Cherry 2015a; Gordon 2011). These scales attempt to identify individuals who may be answering the questions disingenuously which in turn could also lead to providing insights into the psyche of the individual. The majority of the scales in this section are designed to detect deception — whether it be trying to appear better or worse than a person actually is (Cherry 2015a; Gordon 2011). There are 217 items in the validity scales which are mixed in amongst the clinical scales which are all designed to increase the validity of the model and reduce the combat the effects of respondents lying (Cherry 2015a; Gordon 2011; Greene 1990; York 2014).

3.1.2.3 Final Analysis of the MMPI

In summary, the MMPI focus began very clinically and this has been reflected in the uses that the field has found for it and the great effort that was placed on its validity scales also means that sensitive areas such as forensics is able to rely on a personality measure that is hard to fool. However, this focus also means that it is not as applicable in a general setting and context and therefore was excluded from this study.

3.1.3 Factored Model of Personality

The factored models of personality include the HEXACO (or Six Factor) Model and the Five Factor Model (FFM) which are both discussed in (see Section 2.2.1). The HEXACO model was used in fewer studies as compared to the FFM as shown in a review of the literature with the HEXACO model appearing six out of 147 studies and the FFM appearing 79 out of the 147 studies (see Section 2.2.1).

As both models share a substantial portion of their development and purpose, much of the research utilizing both models overlap (see Section 2.2.1). As such, the following sections will discuss both models at once.

3.1.3.1 Area of Use of the Factored Models of Personality

The lexical origin of the factored personality model makes it an ideal tool for the purpose of research between different cultures and ethnicities. By searching for universal aspects of personality across languages, we are also able to develop better understanding of universal traits, values, and behaviours shared by geographically distinct (McCrae et al. 2010; Tackett et al. 2012; Gurven & Rueden 2013).

Since personality itself refers to a relatively stable pattern of behaviour, affect, and thinking (Saucier et al. 2013), it stands to reason that there also exists a fair amount of behavioural research where the 'how' and 'why' of human decisions are picked apart from the perspective of personality (Hilbig et al. 2014; Quilty et al. 2014; Dobewall et al. 2014). This line of research aims to create frameworks that support self-reflection to understand our own decisions and also give us insight into other's decision making to foster understanding between individuals. In furthering that goal, the factored personality models have also found applications in categorizing and understanding nuances in relationships (Furler et al. 2014; Nezlek et al. 2011).

Bridging the gap with health, the factored personality models have had and continue to have a massive impact with its promise of great predictive power (Widiger & Presnall 2013). Specifically, some work has been put into tackling problems associated with depression and anxiety (Lewis et al. 2014; Chow & Roberts 2014), cardiovascular risk (Gleason et al. 2014), and metabolism (Human et al. 2013; Israel et al. 2014) among others. This line of research holds the potential for early detection of complicated health problems that can be mediated by lifestyle changes and save money and lives.

Unsurprisingly, the factored personality models have also been the basis for many contributions to our knowledge of personality disorders (Gleason et al. 2014; Widiger & Costa 2012). Importantly, the FFM has been used heavily in concord with the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5) which is the American Psychiatric Association's (APA) classification and diagnostic tool which is a critical tool in the diagnosis and subsequent treatment of mental problems (Trull 2012; Widiger & Presnall 2013; Mullins-Sweatt & Lengel 2012).

In addition, the factored personality models have been used in several longitudinal studies that aimed to explore the link between early life development and intellect scores (Chamorro-Premuzic et al. 2004; DeYoung et al. 2014; Abe 2005). Those studies showed promise for the use of early personality testing to predict varied factors such as self-regulation and academic performance which can lead to early interventions for poor developmental habits (Chamorro-Premuzic et al. 2004; DeYoung et al. 2014; Abe 2005).

Thus, as is evidenced by the examples in this section, the factored models of personality accepted as the general purpose model for personality in the area of personality research.

3.1.3.2 Validity of the Factored Models of Personality

Examination of the validity of the factored models of personality can be traced all the way back to its conception. Research in 1961 identified recurrent factors of personality that spanned different populations (Tupes & Christal 1961), though the solidification of the factors of personality only occurred several decades later (John et al. 1984; McCrae & John 1992; Goldberg 1971). The formulation of the questionnaires used by this model of personality come from administering a large question pool to a live population that has been professionally diagnosed and then subsequently using factor analysis on the results to form distinct facets of personality (John & S Srivastava 1999; Goldberg 1971). It is based on that work that current personality questionnaires are based on, matching factor and cluster analysis scores with the earlier works in order to prove their validity (Johnson 2014; John & S Srivastava 1999; Costa & McCrae 2008).

3.1.3.3 Final Analysis of the Factored Models of Personality

The factored models of personality (HEXACO and FFM) were designed to capture the general case of personality, focusing on identifying truisms across multiple population samples (Tupes & Christal 1961; John & S Srivastava 1999). They both boast strong validity within the academic community through constantly checking new questionnaire tools against historic mathematic validity measures (Johnson 2014; John & S Srivastava 1999; Costa & McCrae 2008). In comparison of the two models, the FFM is used in more studies than the HEXACO model (as shown in Section 2.2.1) and is therefore chosen as the model of personality to be used as the theoretical basis of the designed video game tool.

3.1.4 Summary of Comparisons Between Personality Models

The most prominent models of personality are the MBTI, MMPI, HEXACO, and FFM. These models normally differentiate themselves around their applied fields, for example, the MBTI is largely used in corporations (Isachsen & Berens 1998), the MMPI in clinical and forensic settings (Pope et al. 2006), and the FFM and HEXACO model in academia (John & Srivastava, 1999). The MBTI suffers from severe criticisms to its validity and the MMPI's primary focus is in clinical psychology which removes them from consideration. The factored models of personality both model general personality, but the FFM is significantly more prominently used over the HEXACO model as shown in the systematic literature review. For those reasons (summarized in Table 5), the FFM is chosen as the model to use in this research.

	Area of Use	Validity
MBTI	General model used in corporate settings	Weak
MMPI	Clinical model used in clinical and forensic settings	Strong
HEXACO	General model used in variety of studies	Strong
FFM	(Most common) General model widely used in variety of studies	Strong

Table 5 – Summary of area of focus of personality models

3.2 Components of the Five Factor Model

The previous section described the process of selecting the FFM as the model of personality to be used as the basis for this project. Contained within each factor of the FFM are six facets. Traditional questionnaire items each correspond to a single facet with the total score among all facets describing a single factor (this breakdown is described in more detail with examples later in Section 4.2). The following sections will describe the five factors of the FFM as well as provide an overview of the corresponding facets. This description includes a trait adjective with each facet which serves to provide a better understanding of the facet being measured. These trait adjectives are typically associated with the NEO PI-R although the primary tool being used in this project is the IPIP (both of which will be fully discussed in Section 3.3).

3.2.1 Extraversion

This trait reflects how much one is oriented towards things outside themselves and includes characteristics such as excitability, sociability, talkativeness, assertiveness, and high amounts of emotional expressiveness (Personality-testing.info 2015; Cherry 2015b). High scorers are said to be extroverts who are energized by socialization, value stimulation, and are usually good at social interaction due to experience (Personality-testing.info 2015). Conversely, low scorers are said to be introverts who tend to be tired out by socialization, value down time, and tend towards the socially awkward (Personality-testing.info 2015). The facets and associated trait adjectives for extraversion are shown in Table 6.

Facet	Trait adjective
Gregariousness	Sociable
Assertiveness	Forceful
Activity Level	Energetic
Excitement-Seeking	Adventurous
Cheerfulness	Enthusiastic
Friendliness	Outgoing

Table 6 – Extraversion facets along with associated trait adjective (John & Srivastava, 1999)

3.2.2 Agreeableness

This trait reflects how much one likes to trying to please others and includes attributes such as trust, altruism, kindness, affection, and other prosocial behaviours (Personality-testing.info 2015; Cherry 2015b). High scorers tend to believe that other people are honest, decent, and trustworthy while low scorers are characterized by scepticism about other people's motives that results in suspicion and unfriendliness (Personality-testing.info 2015). Those very low on agreeableness have a tendency to be manipulative in their social relationships and are more likely to compete than cooperate (Personality-testing.info 2015). The facets and associated trait adjectives for extraversion are shown in Table 7.

Facet	Trait adjective
Trust	Forgiving
Morality	Not demanding
Altruism	Warm
Cooperation	Not stubborn
Modesty	Not show-off
Sympathy	Sympathetic

Table 7 – Agreeableness facets along with associated trait adjectives (John & Srivastava, 1999)

3.2.3 Conscientiousness

This trait reflects how careful and orderly an individual is and is characterized by high levels of thoughtfulness with good impulse control and goal-directed behaviours (Personality-testing.info 2015; Cherry 2015b). High scorers are generally hard working and reliable and at the extreme end may be considered workaholics, perfectionists, and compulsive in their behaviour (Personality-testing.info 2015). Low scorers tend to be more laid back, less goal oriented, less driven by success and are more likely to engage in antisocial and criminal behaviour (Personality-testing.info 2015). The facets and associated trait adjectives for extraversion are shown in Table 8.

Facet	Trait adjective
Self-Efficacy	Efficient
Orderliness	Organized
Dutifulness	Not careless
Achievement Striving	Thorough
Self-Discipline	Not lazy
Cautiousness	Not impulsive

Table 8 – Conscientiousness facets along with associated trait adjectives (John & Srivastava, 1999)

3.2.4 Neuroticism

This trait is the tendency to experience negative emotions and is associated with emotional (in)stability, anxiety, moodiness, irritability, and sadness (Personality-testing.info 2015; Cherry 2015b). High scorers are more susceptible to feelings like anger, envy, guilt, depression and respond poorly to stressors, interpreting ordinary situations as threatening, and minor

frustrations as hopelessly difficult (Personality-testing.info 2015). Neuroticism is a risk factor for the internalizing mental disorders such as phobia, depression, panic disorder, and other anxiety disorders traditionally called neuroses (Personality-testing.info 2015). The facets and associated trait adjectives for extraversion are shown in Table 9.

Facet	Trait adjective
Anxiety	Tense
Anger	Irritable
Depression	Not contented
Self-Consciousness	Shy
Immoderation	Moody
Vulnerability	Not self-confident

Table 9 – Neuroticism facets along with associated trait adjectives (John & Srivastava 1999)

3.2.5 Openness

This trait reflects how much an individual seeks out new experiences and is characterized by features such as imagination and insight (Cherry 2015b; Personality-testing.info 2015). High scorers tend to have a broad range of interests (Cherry 2015b). The facets and associated trait adjectives for extraversion are shown in Table 10.

Facet	Trait adjective
Intellect	Curious
Imagination	Imaginative
Artistic Interests	Artistic
Adventurousness	Wide interests
Emotionality	Excitable
Liberalism	Unconventional

Table 10 – Openness facets along with associated trait adjectives (John & Srivastava, 1999)

3.2.6 Summary of the FFM

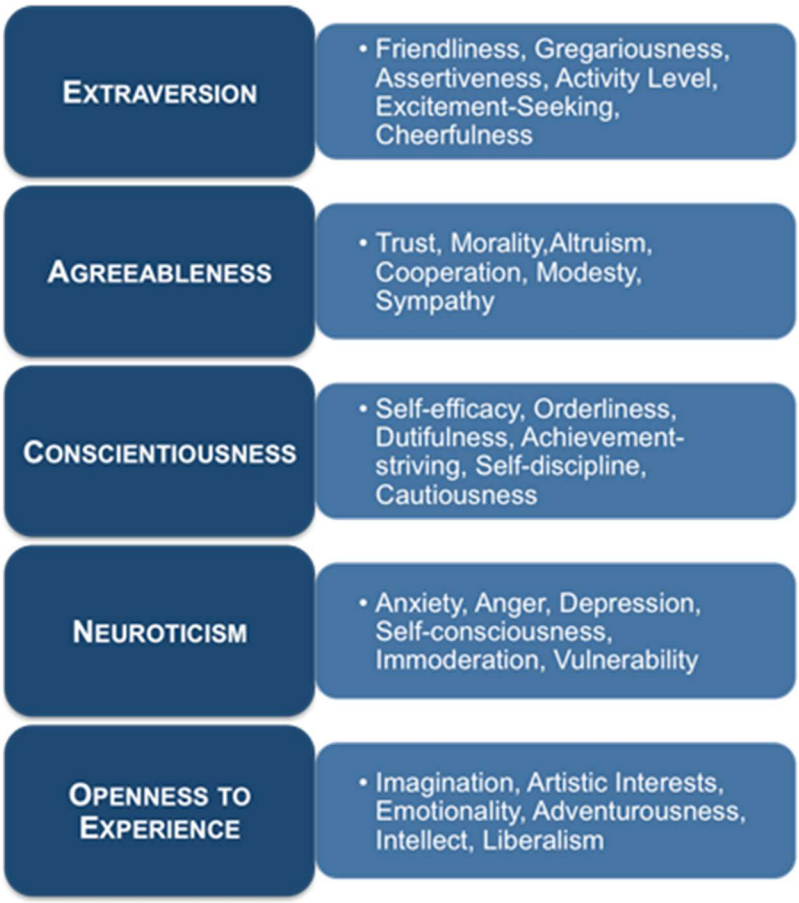

Figure 14 – Summary of Facet breakdowns of each Factor of the FFM

In summary each trait is itself made up of facets and represents a spectrum of human behaviour. Individuals will score anywhere along that spectrum with potentially different scores in each facet which are then combined to provide the Factor score. The facets of each Factor is presented in summary in Figure 14. The general and widespread use of this model (especially in academic research) is why the FFM is chosen as the model to use for this project.

3.3 Questionnaire Selection

Although questionnaires have been shown to have weaknesses in terms of data elicitation, they still serve as the primary method for data collection and acts as the measuring post to which all other methods are compared to. The previous sections described the process of selecting a model of personality as well as gave a brief overview on the components of that model. This section will discuss the selection of a questionnaire from that model to be used as the base for informing decisions in the video game tool's design. This selection is done based

on two criteria: how often a questionnaire is used, cost of obtaining said questionnaire, and how suitable it is for the purposes of the project.

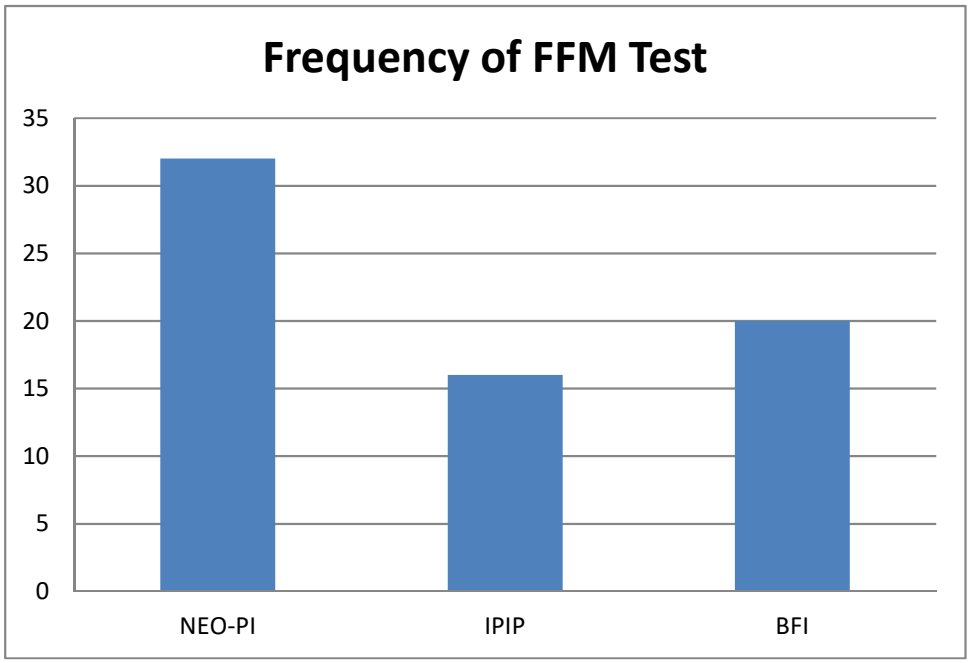

Figure 15 – Frequency of FFM questionnaire occurring in literature reviews

Figure 15 describes the frequency that the most popular questionnaires appeared in the literature review described in Section 2.2.1. The NEO Personality Inventory (NEO-PI) (Costa & McCrae 1985) was used in 32 studies, the International Personality Item Pool (IPIP) (Goldberg 1999) in 16 studies, and the Big Five Inventory (BFI) (John et al. 1991) in 20 studies. The following sections describe those three questionnaires in more detail as well as how those questionnaires meet the criterion identified above.

3.3.1 Big Five Inventory (BFI)

The Big Five Inventory (BFI) was designed to address the need for a short instrument that measures prototypical components of the FFM (John et al. 1991). It has 44 items and was developed through expert ratings and subsequent factor analytic verification in observer personality ratings (John & S Srivastava 1999; John et al. 1991). In order to create such a brief questionnaire, the BFI ignores facet level information and focuses on reporting at the factor level (John & S Srivastava 1999; John et al. 1991). Given the scope of the project, it was determined that the granularity afforded by obtaining facet level information would be important to the design process. It is for that reason that this inventory was not chosen as the questionnaire to be used as the basis for this project.

3.3.2 NEO Personality Inventory (NEO-PI)

The NEO Personality Inventory (NEO-PI) was developed in samples of middle-aged and older adults, using both factor analytic and multi method validational procedures of test construction (Costa & McCrae 1985; Costa & McCrae 1992). The scales have shown substantial internal consistency, temporal stability, and convergent, and discriminant validity against spouse and peer ratings (John & S Srivastava 1999; Costa & McCrae 2008). However, while it acts as the gold standard for personality testing, it is a proprietary tool and thus costs a substantial amount of funds as well as requiring the purchaser to prove that they are qualified to administer the questionnaire (Costa & McCrae 2016). Thus, for the purposes of this study, the NEO-PI was not chosen as the questionnaire to be used as the basis for this project.

3.3.3 International Personality Item Pool (IPIP)

The International Personality Item Pool (IPIP) was created as a public-domain resource to be used as a reliable and valid alternative to commercial inventories like the NEO-PI (Goldberg 1999). One of the first personality measures to be created from the IPIP was a 300-item Inventory designed to measure constructs similar to those assessed by the 30 facet scales in the NEO PI-R (Goldberg 1999).

A revised version of the 300-item inventory called the IPIP-NEO was designed and developed with the capability to be administered on the World Wide Web (Johnson 2008). The first published studies with the IPIP-NEO indicated that the scales of this inventory showed a mean alpha reliability of 0.80, surpassing the mean alpha of 0.75 for the original NEO PI-R scales (Johnson 2008). In response to the length of these questionnaires, a shortened 120-item scale was developed and validated against an internet sample of N = 21,588 and was called the IPIP NEO-PI-R which covers the five factors and six facets per factor of personality, representing each facet with four questions (Johnson 2014). That questionnaire (along with all other IPIP variants) can be found online with instructions on administration in an open source format (Software 2014). The IPIP NEO-PI-R was selected to be the basis for the design process for the video game tool of the project because it is readily available and accessible, provides a substantial amount of detail in the way the question items are related to personality facets, and is a popular and regularly used questionnaire within the area of personality research.

3.3.4 Summary of the Questionnaire Selection

In order to select a questionnaire to be referenced during the design process for the video game tool, three popular questionnaires were identified and analyzed. The BFI was rejected due to its lack of granularity as it was designed to be brief and report only at the factor level. The NEO-PI was rejected due to the cost associated with obtaining the questionnaire which not only covered monetary expenses, but also technical expertise. The IPIP was ultimately chosen because it provided instruction on its use and interpretation in an open source manner as well as offering questions which correspond to specific facets of the FFM which will be useful in the design phase of the project.

3.4 Deriving Game Design from a Questionnaire

The previous sections described the selection of a personality model as well as a questionnaire to be used as the basis for the development of a video game personality elicitation tool. That served as the first step to answering SQ1 as defined in this chapter. This section will discuss the process of taking items from a questionnaire and creating a game design specification.

3.4.1 Questionnaire Item to Game Design Process

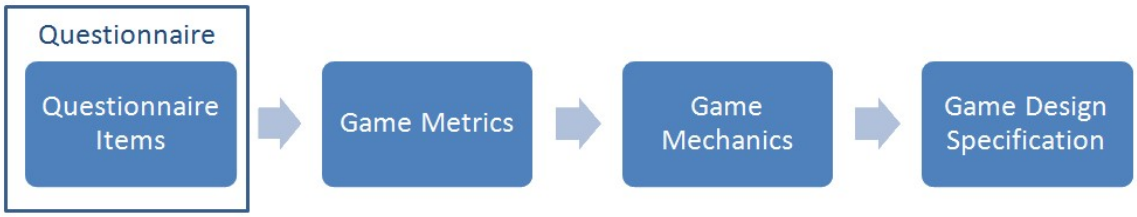

Figure 16 – Game design derivation process summary

In order to develop a game design that best suited the task of personality data elicitation, a rigorous design process (summarized in Figure 16) was undertaken in order to identify the behaviours that best captures personality data to be measured by the video game tool. This process consisted of four main steps:

1. Consultation with subject matter experts from both the areas of personality and game design to identify questionnaire items and associated behaviour to be measured.
2. The identification of game metrics that best captured the identified behaviours.
    a. Compiling the best and most feasible ideas into an exhaustive list of game metrics that could be used for a given facet/sample question.

3. The discussion on which game mechanics are needed in order to implement the identified game metrics while adhering to principles of good game design (Jesse Schell 2008; Despain et al. 2013; Salmond 2016).
    a. The debate on how these game mechanics best fit together into a cohesive game and then listing the possibilities for combinations of these mechanics.
4. The final decision on which game mechanic to be placed into the video game tool.

The process begins with step 1: an examination of the full 120 questions of the IPIP NEO-PI-R (Johnson 2014). From there, subject matter experts identify questionnaire items that can best be measured by in-game metrics through discussion and relying on the experience of the experts from the area of personality to identify the underlying behaviour to be measured.

This process of deriving the proposed game's design from the questionnaire was done in consultation with subject matter experts in the area of personality research. Professor Ken Kirkby, a professor of Psychiatry, and Dr. Allison Matthews from the division of Psychology within the School of Medicine and both from the University of Tasmania were brought on board the project to facilitate this role.

Step 2 begins a process of discussion over the nature of the behaviour identified in step 1 as well as how that kind of behaviour can best be observed in a video game environment. This step seeks to identify as many possible ideas for metrics as possible for a given behaviour through discussion with subject matter experts. For example, for the sample question "Jump into things without thinking" from the facet of Cautiousness under the factor of Conscientiousness, a list of promising possible metrics were:

1. How the player approached a puzzle;
2. How often the player used the map function;
3. How long the player spent deliberating on moves during their turn; and
4. How often the player paused the game.

From there, step 3 began discussing what mechanics were needed in a game in order to measure the metrics identified in step 2. In the example of the list given above, the game would need the capacity for either:

1. A puzzle that could be solved from first glance;
2. A map function in a game where viewing the map can be useful for progression;
3. A turn-based gameplay system that required the player to react to events that happened out of their turn; or
4. A game where pausing the game can provide some form of advantage for the player.

Step 3 also extended into discussing how game mechanics identified here could best be combined to create a cohesive game experience. This was done by relying on expertise from the video game design experts, but also on literature for good design practices (Jesse Schell 2008; Despain et al. 2013; Salmond 2016). Some of the initial ideas developed at this point at the early points of this design phase are:

1. A top-down exploration game with limited use weapons that are held within an inventory. This game would have standard RPG mechanics with a quest system.
2. An inventory game where the goal was to manage an inventory system as items were piped into the player's play area with requests for items coming in every now and then.
3. A side-scrolling platforming game where the player has to use items within their inventory to traverse the world towards a goal or exit.

Finally, step 4 was a finalization phase where all of the previous steps' discussion culminated in the addition of a new mechanic added to the video game tool. This was always done in consultation with the subject matter experts as a final review to be clear that the assumptions that had been made along the way were well founded.

In the case of the initial game ideas shown above, the top-down exploration game would have been relatively simple to design for as it played to many common clichés in video games. However, it was ultimately discarded as a potential game idea due to being extremely expensive in terms of time and resources needed to create a compelling experience within that game archetype.

The puzzle inventory game idea was rejected due to the immediate apparent complexity associated with designing puzzles that would be needed for such a game. It was likely that the best tools for such a game would require some element of time pressure which was ruled

in order to create a video game tool that would be as accessible as possible in an attempt to avoid alienating any potential players that may arise from a game that requires fast reactions.

The side-scrolling platforming game idea utilizes relatively simple mechanics, relying primarily on movement and resource management. The goal of the game is straightforward (find the exit), and this game concept seemed to be something that could be prototyped quickly to allow for more complexity and depth to be added as the discussion process continued and more game mechanics were identified. Thus, the side-scrolling platforming game idea was chosen as the basis for the video game tool to be designed.

Once the game idea was agreed upon and locked down, the 4-step process was repeated until the team of subject matter experts could no longer find any more sensible behaviours to be converted to game mechanics. In these cases, extra emphasis was also placed on how the newly identified game mechanics could be integrated as seamlessly as possible into the existing game design, creating a process that grew a game out from a small set of game mechanics by iteratively adding on new elements.

The full list of game mechanics added to the game as well as the sample behaviour it is attempting to measure and the rationale for that addition is described in the following sections and is summarized in Section 3.4.3.

3.4.2 Process of Converting Questionnaire Items to Game Mechanics

In order to derive game metrics from the questionnaire, sample questions were identified to be used as the basis for design of the video game tool. These sample questions were then used to identify a behaviour that could be measured within a video game context – a metric within the game. Finally, these metrics were used to identify the necessary game mechanic to allow for that behaviour to be measured.

3.4.2.1 Conscientiousness

Six sample questions were identified in the Conscientiousness trait of the FFM with one corresponding to each facet.

3.4.2.1.1 Orderliness

The questions associated with the facet of Orderliness are: "Like to tidy up", "Often forget to put things back in their proper place", "Leave a mess in my room", and "Leave my belongings

around" with the latter three being negatively keyed. All items in associated with this facet revolve around keeping items and belongings in a state of sorder.

Experts from the field of personality described this as a means for the individual to exert their order upon an environment. Experts from the field of game design proposed that this was best accomplished with management of some form, whether it be physical in which the player is required to be aware of objects in the game world as they traverse it, or a more abstract sense where the player is keeping track of the order of items that are available to them. Thus, the potential game would require some manner in which the player may exert their order upon the game world.

It was decided that an inventory system with a series of different items which is a common video game mechanic would serve this purpose (Moore 2011). It was hypothesised by experts in both fields that high scorers for the question of "Like to tidy up" would choose to put more effort into keeping their inventory ordered. The actual logic behind ordering this inventory is not as important as the sheer act of putting time and effort into the organization in the first place. This element required not just the design of an inventory system, but also the creation of a game that uses items as part of its mechanism.

This game mechanic became one of the first ideas explored for the game. After the initial analysis of the items within the questionnaire, this set of questions were one of the first that strongly lent themselves to a task that could be done within a video game environment. This early identification of an inventory and item system critically influenced the subsequent design of the video game and serves as one of a number of factors that pushed the game away from an action-reaction based game.

3.4.2.1.2 Self-Efficacy

The questions associated with the facet of Self-Efficacy are: "Complete tasks successfully", "Excel in what I do", "Handle tasks smoothly", and "Know how to get things done" where all items are positively keyed. All items associated with this facet revolve around the concept of performing tasks in a general sense with a high amount of competency.

Experts from the field of personality described this as an individual's inherent competence at completing tasks being important to themselves. Game design experts offered that the video

game would provide a rating of how well the player has performed which would allow them to internally understand whether or not they are performing well and subsequently adjust the way they player (whether it be a conscious decision or not). Traditionally this has been done in numerous methods such as star-ratings, percentage completion ratios, or simply just score ratings serving as common ways to track performance in video games (von Ahn & Dabbish 2008).

However once the game design was finalized, it became apparent that the primary goal of the game was to complete the given level, with extra score being rewarded for performing additional collection tasks. Therefore, the game metric for the question "Excel in what I do" was revised to be the time taken for the player to complete the level as that was a more direct measure of how well the task was completed. It is hypothesised that high scorers in the question would seek to complete the task as swiftly as possible.

While this facet seems an obvious and simple measure on the surface, the actual measure required for this facet can change drastically with the design of the game. Small tweaks to an existing design could alter the main goal of the game. For example, by shifting focus into exploration and rewarding the player as they discover more areas in a game level would completely change the goal of the designed video game while keeping large portions of the gameplay intact.

3.4.2.1.3 Cautiousness

The questions associated with the facet of Cautiousness are: "Jump into things without thinking", "Make rash decisions", "Rush into things", and "Act without thinking" where all items are negatively keyed. All items associated with this facet revolve around the amount of deliberation a person may take when dealing with tasks in their everyday life primarily in situations that are not immediately familiar and outside of their routine.

Experts in the field of personality describe this measure as the degree to which an individual may act spontaneously, or conversely how much deliberation they choose to take when confronted with something new. Game design experts offered a scenario where a player is confronted with new information such as seeing a level or a puzzle for the first time and then measuring the time they spend before tackling that problem. This plays into common problem

solving skills found in many video games (Moore 2011) and the amount of time spent in deliberation before action during those periods will provide insight into this facet of their personality. Thus, the potential game would require some aspect that would cause the player to pause and consider the options that are available to them

It was decided that the primary means of measuring this facet would be the amount of time the player spends before starting a puzzle. While the player's first interaction with a level can also yield interesting results, a one-screen puzzle can be placed wholly in the player's view easily, giving them all the information necessary to solve it. Whether the player chooses to take their time to work out a solution before attempting to solve the puzzle becomes an indication of how deliberate they are. It is hypothesised that players who score highly in the question "Jump into things without thinking" (and therefore lowly in the facet due to a negative correlation) will prefer to approach the puzzle by immediately testing moves and hoping to stumble across a solution as opposed to solving the puzzle in their head and transferring that to the game.

This mechanic initially focused on the player's desire to use the map or tutorial functions in the game. However, as other facets as described in Sections 3.4.2.3.1, 3.4.2.3.2, and 3.4.2.3.3 made the addition of a traditional one-screen puzzle sensible to the game, it was determined that this metric was better measured in the context of a puzzle. The major point of differentiation is how the player is able to take in all of the information at once since the puzzle (and solution) all exist on one screen.

3.4.2.1.4 Achievement Striving

The questions associated with the facet of Achievement Striving are: "Do more than what's expected of me", "Work hard", "Put little time and effort into my work", and "Do just enough work to get by" with the latter two being negatively keyed. All items associated with this facet revolve around the concept of exceeding the expectations of a task presented to the individual.

Experts from the field of personality describe this measure as the desire for an individual to be seen as someone who has done more than was asked of them. Game design experts associated this behaviour with the concept of tertiary goals such as side quests or collection tasks which do not impede the completion of primary goals, but provide an opportunity for

players to accomplish more than just the bare minimum in a game (Moore 2011). This is further solidified by the presence of a leaderboard for scores which act as a way to rank a player's performance (Seaborn & Fels 2015; Deterding et al. 2011). Thus, the potential game requires some means by which the player may assess their performance in a task in order to decide if they would like to improve upon their old performance.

It was agreed upon by both experts in both the fields of personality and game design that a collection mechanic is the simplest implementation of the tertiary goal described above. Due to this, collectable coins were added to the game with each coin contributing to the player's total score in the level.

3.4.2.1.5 Self-discipline

The questions associated with the facet of Self-Discipline are: "Am always prepared", "Carry out my plans", "Waste my time", and "Have difficulty starting tasks" with the latter two being negatively keyed. All items associated with this facet revolve around the concept of planning.

Experts in the fields of personality describe this measure as the degree to which an individual would care about planning out actions in anticipation of upcoming challenges. Game design experts targeted this measure by measuring the amount of time players are willing to spend collecting resources (in this case items such as ladders, bridges, and ropes) that can be used to overcome obstacles later on in the game. Thus, the game offers the player an ability to make preparations that will benefit them later on in the level.

The simplest measure of this metric was simply to consider how many times the player interacted with items and their inventory within a level. It is hypothesised that players who score highly in the question "Am always prepared" will spend more time and effort to prepare for possible (and unknown) obstacles ahead of time instead of simply moving on to the next task.

The discussions held between experts during the design phase for the game aimed to target the game to as broad of a representation of the personality inventory as possible. In some cases, the lines between where one metric ended and another begins is blurred such as the case between this facet and Deliberation (see Section 3.4.2.1.3). In these cases, the underlying personality inventory has the mathematical clustering to clearly delineate between these

items. But the design of this game is not as neat and oftentimes the effort to include as many measures as possible (so as to potentially find connections) skirts close to blurring two metrics together. It is hypothesised that some clear delineation between these metrics could be discovered once the game can be tested at a large enough scale, but for now remains merely a point of note in the design of the game as something worth including on the probable link to a personality facet.

3.4.2.1.6 Dutifulness

The questions associated with the facet of Dutifulness are: "Keep my promises", "Tell the truth", "Break the rules", and "Break my promises" with the latter two being negatively keyed. All items associated with this facet revolve around the concept of obeying rules.

Experts in the field of personality describe this measure as the extent an individual values obeying the rules. Game design experts determined that this can be measured by the implementation of an intentional 'bug' or mistake in the game that can be exploited by the player for their own advantage. Thus, the game has to offer the player a chance to break the rules as far as they understand the rules to exist within the game.

However, this brought the actual form of the bug itself into question. It needed to be a phenomenon that occurred with relative certainty for every player as opposed to something that required a complicated set up. Further, the rules of a game world are flexible to begin with, so the bug had to be clearly something that was unintended as far as the world's rules go. Lastly, the bug also had to be clearly measurable by the game so that instances where the player chooses to use it can be recorded.

It was ultimately decided that a bug would be introduced into the leaderboard system where the player can alter the score that is submitted to the game. The player's reported score would be placed in a text box as opposed to a static label. Care was taken to make it obvious to the player that the field where the number was placed could be modified. It is hypothesised that players who score highly in the question "break rules" (and lowly in the facet due to negative correlation) will be more willing to exploit a bug in a game to their advantage. Further aspects of this design are discussed in Section 3.4.2.3.4.

3.4.2.2 Openness

Four sample questions were identified in the Openness to Experience trait of the FFM with no sample questions identified for the facets of Emotionality and Liberalism. This omission is discussed in a later section.

3.4.2.2.1 Intellect

The questions associated with the facet of Intellect are: "Love to read challenging material", "Avoid philosophical discussions", "Have difficulty understanding abstract ideas", and "Am not interested in theoretical discussion" with the latter three being negatively keyed. All items associated with this facet revolve around the concept of spending mental energy on thought challenges.

Experts in the field of personality describe this measure as the amount to which the individual would seek out mentally stimulating challenges. Game design experts translated this into the need for a cerebral task for the player where they are able to control the degree of challenge in the task. It was decided after deliberation to include a simple puzzle which also featured different challenge levels which allowed the player to choose the difficulty they desired. At the time, this decision added a substantial new mechanic to the game's design, but it was deemed necessary by experts from both fields as it was important to make it clear to the player that this decision altered the level of challenge they would receive.

In its initial state, the game only consisted of the level traversal challenge with associated item and inventory system. The added puzzle would have the player select a difficulty level with the hypothesis that players who score highly in the question "love to read challenging material" will prefer more challenging puzzles. This may be due to several reasons such as deriving satisfaction from the completion of a more difficult task. For a more comprehensive description of the puzzle that was added, please refer to Sections 3.4.2.3.1, 3.4.2.3.2, and 3.4.2.3.3. Starting from easy difficulty going to medium and finally hard, the player's choice increases the amount of complexity present in the puzzles by expanding the board and adding more blocks that had to be matched to a goal point.

This puzzle was further integrated into the loose narrative of the game to mitigate the feeling of the puzzle being tacked onto a complete game concept. The player is informed that their

exit from the level is sealed away behind the solution of this puzzle in an attempt to build a story around the addition of the puzzle element.

3.4.2.2.2 Adventurousness

The questions associated with the facet of Adventurousness are: "Prefer variety to routine", "Prefer to stick with things that I know", "Dislike changes", and "Am attached to conventional ways" with the latter three items being negatively keyed. All items associated with this facet revolve around the concept of new experiences.

Experts from the field of personality describe this as the desire for individuals to seek out new experiences. Experts from both personality and game design discussed this variety in experiences in context of the variety of gameplay offered by the game, but agreed that since the player isn't making a decision on the kind of variability or when they experience that variability in the game, that the gameplay itself being diverse did not cover this player behaviour.

Instead, game design experts decided that the variety in question has to be an active choice by the player. This could typically be accomplished with variety in gameplay styles. For example, in shooting games, this variety could come from the choice of gun the player chooses. Ultimately, this is best measured when choices made by the player affect the gameplay experience they receive in an orthogonal way such that there is no clear superior option, but that each option provides a different kind of experience.

Unfortunately, when it comes to game designed for this project, there is no available design space for this kind of choice. Attempting to add space for the player to make meaningful gameplay choices would have bloated the design of the game far beyond what is achievable in the scope of this project.

In order to accommodate this metric, options that don't alter gameplay were considered. This consideration happened alongside the discussion over design additions for Artistic Interests and Imagination facets (see Sections 3.4.2.2.3 and 3.4.2.2.4). Due to the fact that a new game element was added for those measures, a design space was opened up to also allow for this facet. It is hypothesised that players who score highly in the question "prefer variety to routine" will elect to alter the look of either the player avatar or the theme of the world.

While not an ideal measure for this facet, the fact that design space was created incidentally through the design process of other facets meant that it was still worth considering how player choices in this aspect of the game reflected in their personality.

3.4.2.2.3 Artistic Interests

The questions associated with the facet of Artistic Interests are: "Believe in the importance of art", "See beauty in things that others might not notice", "Do not like poetry", and "Do not enjoy going to art museums" with the latter two items being negatively keyed. All items associated with this facet revolve around how much the individual values aesthetics.

Experts from the field of personality describe this measure as the individual's own values on aesthetics – not necessarily art in the traditional sense, but an appreciation for artistic diversity. Game design experts posed that an appropriate metric for this measure is the allowance for the player to adjust the graphical aspect of the video game. Thus, the game would allow for the player to make decisions on the aesthetics of the game and therefore allow their exploration of the game's offered artistic styles to be captured.

This can be achieved by changing only the visual appearance of sprites in the game which results in completely different visual styles while leaving the mechanical functionality of the game unchanged. Although ultimately aiming to measure different aspects of player choice, this measure is related to the facet of Imagination as well (see Section 3.4.2.2.4). Thus while the description and justification for what kinds of aesthetic choices are given to the player is discussed in Section 3.4.2.2.4, this game measure is concerned with how often aesthetic changes are made. It is hypothesised that players who score highly in the question "believe in the importance of art" will change the aesthetic look of the game more often.

3.4.2.2.4 Imagination

The questions associated with the facet of Imagination are: "Have a vivid imagination", "Enjoy wild flights of fantasy", "Love to daydream", and "Like to get lost in thought" where all items are positively keyed. All items associated with this facet revolve around valuing elements that are extraordinary.

Experts from the field of personality describe this measure as the individual's value on fantastic and extraordinary elements. Game design experts took the idea discussed in Section

3.4.2.2.3 and built a design which emphasised fantasy graphical components in order to allow players the ability to explore and preferentially choose extraordinary aesthetics.

It is hypothesised players who score highly in the question "enjoy wild flights of fantasy" will gravitate more towards imagery that is fantastical in nature. To that end, alternative sprites for the player character that resembles an orc, a dragon, and an alien were added to the game. These do not functionally change the way the game is played in any way, but rather how the game looks.

3.4.2.3 Agreeableness

Four sample questions were identified in the Agreeableness trait of the FFM with no sample questions identified for the facets of Cooperation and Sympathy. This omission is discussed in a later section.

3.4.2.3.1 Trust

The questions associated with the facet of Trust are: "Trust others", "Believe that others have good intentions", "Trust what people say", and "Distrust people" with the last item being negatively keyed. All items associated with this facet revolve around the individual placing their trust in others.

Experts from the field of personality describe this measure as the individual's trust in anonymous strangers. This and several other facets (see Sections 3.4.2.3.2 and 3.4.2.3.3) presented an interesting challenge to the design. The project lacked any scope to create a social experience for players, but facets such as these deal with the interrelationships of people – how they view a stranger's intentions in this case. Experts in both the fields of personality and game design agree that it would be sufficient to fake the social interactions that are given to players.

In order to obtain a measure for this facet, a gifting system was implemented into the game. Once encountered, players are given the option of accepting or rejecting the gift. If rejected, the player would simply move on and receive no benefits or detriments from the gift. However, if the player accepts it, there is a chance built into the game that either adds to or subtracts from the player's score. The player is told that the gift has been left there by another

player, but in reality the system is using a random number generator to determine the outcome of the gift and is faking the social aspect of this interaction.

Ultimately the choice of the player is the important aspect to be measured. It is hypothesised that players who score highly in the question "trust others" will be more willing to trust strangers in this scenario across repeated interactions. This further serves to potentially increase engagement and investment by the player since it adds an aspect of social depth to the game (Moore 2011).

3.4.2.3.2 Altruism

The questions associated with the facet of Altruism are: "Am concerned about others", "Love to help others", "Am indifferent to the feelings of others", and "Take no time for others" with the latter two items being negatively keyed. All items associated with this facet revolve around the concept of an individual helping another.

Experts from the field of personality describe this as an individual's desire to help anonymous strangers. This and several other facets (see Sections 3.4.2.3.1 and 3.4.2.3.3) presented an interesting challenge to the design. The project lacked any scope to create a social experience for players, but facets such as these deal with the interrelationships of people – whether an individual seeks to help a strange in this case. Experts in both the fields of personality and game design agree that it would be sufficient to fake the social interactions that are given to players.

In order to obtain a measure for this facet, a gifting system was implemented into the game. At the end of a level, the player is asked if they would like to give a positive gift or a negative gift to a stranger playing the game. They may have already come across such gifts during their time in the level as described in Section 3.4.2.3.1 which helps to reinforce the idea that they are able to affect another person's game experience. In reality neither option will cause any tangible difference in the game experience, but the pretext of supplying aid serves as a means to measure the player's desire to help strangers.

It is hypothesised that players who score highly in the question "love to help others" will choose to give positive gifts to other players.

3.4.2.3.3 Morality

The questions associated with the facet of Morality are: "Love a good fight", "Yell at people", "Insult people", and "Get back at others" with all items being negatively keyed. All items associated with this facet revolve around the concept of conflict and generating negative experiences for others.

Experts from the field of personality describe this as an individual's desire to cause conflict or negative experiences for others. This and several other facets (see Sections 3.4.2.3.1 and 3.4.2.3.2) presented an interesting challenge to the design. The project lacked any scope to create a social experience for players, but facets such as these deal with the interrelationships of people – whether an individual seeks to negatively impact another's experience in this case. Experts in both the fields of personality and game design agree that it would be sufficient to fake the social interactions that are given to players.

In order to obtain a measure for this facet, a gifting system was implemented into the game. At the end of a level, the player is asked if they would like to give a positive gift or a negative gift to a stranger playing the game. They may have already come across such gifts during their time in the level as described in Section 3.4.2.3.1 which helps to reinforce the idea that they are able to affect another person's game experience. In reality neither option will cause any tangible difference in the game experience, but the pretext of causing harm serves as a means to measure the player's desire to create conflict.

It is hypothesised that players who score highly in the question "get back at others" will choose to give negative gifts to other players.

3.4.2.3.4 Modesty

The questions associated with the facet Modesty are: "Believe that I am better than others", "Think highly of myself", "Have a high opinion of myself", and "Boast about my virtues" with all items being negatively keyed. All items associated with this facet revolve around the concept of modesty in their own individual abilities.

Experts in the field of personality describe this measure as an individual's desire to show off. Game design experts likened this concept to that of a leaderboard system. However, the nature of the game and a potentially small sample size of players would mean that most

leaderboards would be underpopulated. This is another instance where the social context is important, but can be faked.

At the end of each level, the player is presented with a list of scores that showcase other players' efforts in that same level – a leaderboard. This leaderboard is populated with random values to imply that many individuals have played the level while in reality those numbers are generated to fall within the range of possible scores for that level. Players are then given the option of uploading the score that they had obtained in the level to the perceived community leaderboard. It is hypothesised that players who score highly in the question "think highly of myself" (and lowly in the facet due to a negative correlation with the question score) will be more inclined showcase their achievement in this opportunity. This game element also ties into Dutifulness (see Section 3.4.2.1.6).

3.4.2.4 Extraversion

Two sample questions were identified in the Extraversion trait of the FFM with no sample questions identified for the facets of Excitement-seeking, Friendliness, Gregariousness, and Activity Level.

3.4.2.4.1 Cheerfulness

The questions associated with the facet Cheerfulness are: "Radiate joy", "Have a lot of fun", "Love life", and "Look at the bright side of life" with all items being positively keyed. All items associated with this facet revolve around the concept of finding enjoyment in mundane aspects of everyday routine.

Experts in the field of personality describe this measure as an individual's desire to find and create happiness in others. The game has to give the player an opportunity for this tendency to shine through in their gameplay decisions and ideally this would be done by allowing interactions between actual human players. However, the addition of mechanics that would allow truly social interactions in the game along with fully testing, integrating, and ensuring functional integrity of corresponding game elements would vastly exceed the scope of the project.

The existing game design is asynchronous in nature and not dependent on another individual to provide the same level of experience. Any effort to add truly social features would

introduce numerous 'first-time' edge case scenarios, lack of other player data scenarios, and generally make the game experience more dependent on the size of its player base and compromise its ability to deliver a consistent single player experience.

Previously described similar mechanics that required social interaction were faked (see Sections 3.4.2.3.1, 3.4.2.3.2, and 3.4.2.3.3), but there was no achievable interaction that could be faked for this measure. After much discussion, experts in both the fields of personality and game design agree that non-playable characters (NPCs) would serve as adequate proxies for the purposes of allowing the player to express themselves in this facet. NPCs were thus added to each level. These NPCs do not serve to aid or impede player progress. While the player is within a short proximity with an NPC, they are able to push buttons that will display an emoticon of either a smiling, frowning, crying, or angry face over their head. This will be reflected in every NPC nearby.

It is hypothesised that players who score highly in the question "radiate joy" will choose to interact in a socially positive way given this opportunity. While this is not a perfect system for this facet, it was deemed acceptable as a solution to attempt to cover the factors of the FFM of personality as widely as possible.

3.4.2.4.2 Gregariousness

The questions associated with the facet of Gregariousness are: "Love large parties", "Talk to a lot of different people at parties", "Prefer to be alone", and "Avoid crowds" with the latter two being negatively keyed. All items associated with this facet revolve around the individual seeking to surround themselves with other people.

Experts in the field of personality describe this measure as an individual's desire to be surrounded by others. Thus, in order to take measures that relate to this tendency, the game has to offer the player opportunity to seek out social gatherings. Similar to the Cheerfulness facet (see Section 3.4.2.4.1), this is a highly social phenomenon that is difficult to replicate in such a strictly single player experience. However, experts from both the fields of personality and game design agree that with the addition of an NPC system, the game may sensibly take a measure of how often the NPCs are sought out by the player as an indication for this facet.

In order to put the choice in the player's hands, NPCs within each level are deliberately placed within view of the player, but out of their optimal path to the exit. In doing so, any time the player is recorded as having sought out the NPCs, it is done consciously. It is hypothesised that players who score highly in the question "avoid crowds" (and therefore lowly in the facet due to negative correlation) will choose to avoid areas with NPC clustering if possible.

3.4.3 Summary of Questionnaire Item to Game Mechanics Conversion

The previous sections described the rationale taken when converting questionnaire items to game metrics and finally to game mechanics. A summary of the outputs of that process is provided in Table 11.

Trait	Facet	Sample question	Game metric	Game element
Conscientiousness	Orderliness	Like to tidy up.	Orderliness of a player's inventory	Inventory system
	Self-Efficacy	Excel in what I do.	How well a player performs and whether or not they try to improve their score	Scoring system that showcases player performance
	Cautiousness	Jump into things without thinking.	Amount of time a player spends deliberating a puzzle	Predictable, observable, and deterministic puzzles
	Achievement Striving	Do more than what's expected of me.	How many side quests a player completes	Side quests
	Self-discipline	Am always prepared.	How meticulous a player is in anticipation of future obstacles	Allow preparation in anticipation of obstacles
	Dutifulness	Break rules.	How often a player breaks the rules of the game	Glitches or abuses that are intentionally left in the game
Openness	Intellect	Love to read challenging material.	Choosing difficulty challenge	Puzzles with different complexities
	Adventurousness	Prefer variety to routine.	How often a player changes play styles of game elements	Choices in gameplay or aesthetics
	Artistic Interests	Believe in the importance of art.	Whether a player spends time and effort on customising the game aesthetic	Choices in aesthetics
	Imagination	Enjoy wild flights of fantasy.	Whether a player prefers fantastical game aesthetic to mundane ones	Fantastical and mundane aesthetics
Agreeableness	Trust	Trust others.	Whether a player trusts gifts that can be positive or negative from other players	Between-player gift interaction
	Altruism	Love to help others.	Whether a player chooses to give a positive or negative gift to other players	Between-player gift interaction
	Morality	Get back at others.	Whether a player chooses to retaliate when given a negative gift	Between-player gift interaction
	Modesty	Think highly of myself.	Whether a player chooses to have their name displayed on a leaderboard	Leaderboard system
Extraversion	Excitement-seeking	Enjoy being reckless.	Whether a player takes a gamble situation	Random number generator gamble situation
	Cheerfulness	Radiate Joy	Whether a player chooses to interact positively with others	Simple mood social system
	Gregariousness	Avoid crowds.	Whether a player chooses to go towards others	Branching paths

Table 11 – Questionnaire item to game mechanic conversion summary

3.5 Game Implementation

The previous section described the design process of taking questionnaire items and producing a game design description in order to answer SQ1. This section discusses the implementation of the game design description in order to begin answering SQ2.

3.5.1 Inventory System

An inventory and related item system was implemented in order to accommodate the metrics identified in Sections 3.4.2.1.1 and 3.4.2.1.5. As discussed in Section 3.4.2.1.5, it is hypothesized that players who score higher in the Self-Discipline facet would spend more time preparing for future obstacles by collecting more resources more often. Ropes, ladders, and bridges were implemented into the game as items in the game world that can be gained and placed into the inventory system for this purpose Figure 17.

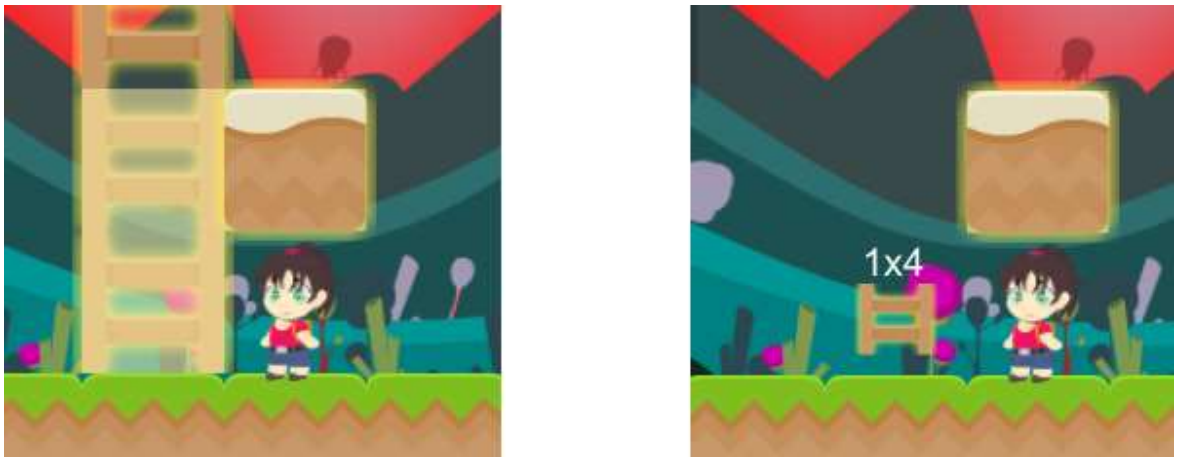

Figure 17- Example of converting a world item to a pickup

While those items occur naturally, players may choose to interact with them as they traverse the level and collect them in preparation for future challenges where they may use them. Each item will occupy a different amount of space in the player's inventory with ropes, ladders, or bridges that are longer or taller taking up more space than their shorter counterparts. For example, a ladder that is two units of world space high will occupy two units of inventory space while a four-unit high ladder will occupy four units of inventory space Figure 18. This forces the player

to choose between carrying fewer but larger items to cover greater distances or more small items to overcome more small obstacles.

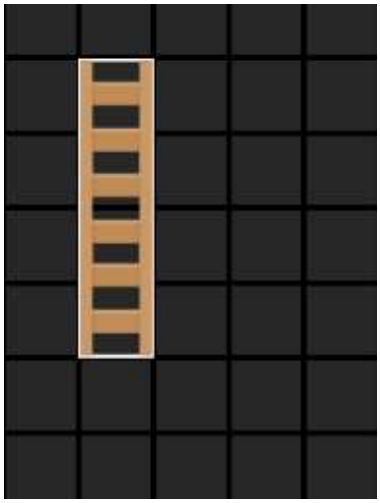

Figure 18 – Ladder item taking up 4 spaces in a vertical shape in the inventory

As discussed in Section 3.4.2.1.1, it is hypothesized that players who score higher in the Orderliness facet would spend time organizing their inventory. In order to encourage this behaviour, the inventory was limited in size to a space that accommodated 35 units worth of items. This increases the frequency at which players will have to interact with the inventory system to manage their resources. It is theorized that high scorers in the Orderliness facet would spend extra time organizing the inventory at this point.

3.5.2 Scoring System

A scoring and related collectible coin system was implemented in order to accommodate the metric identified in Section 3.4.2.1.4. As discussed in Section 3.4.2.1.4, it is hypothesized that players who score highly in the Achievement Striving facet would strive to perform better and therefore seek to obtain more score. Collectible coins (similar to many other score enhancing collectibles in other video games (Despain et al. 2013; Salmond 2016)) were added to the game to facilitate this. The coins are littered around the level for the player to collect or ignore and there is no penalty for skipping the coin collection altogether (Figure 19).

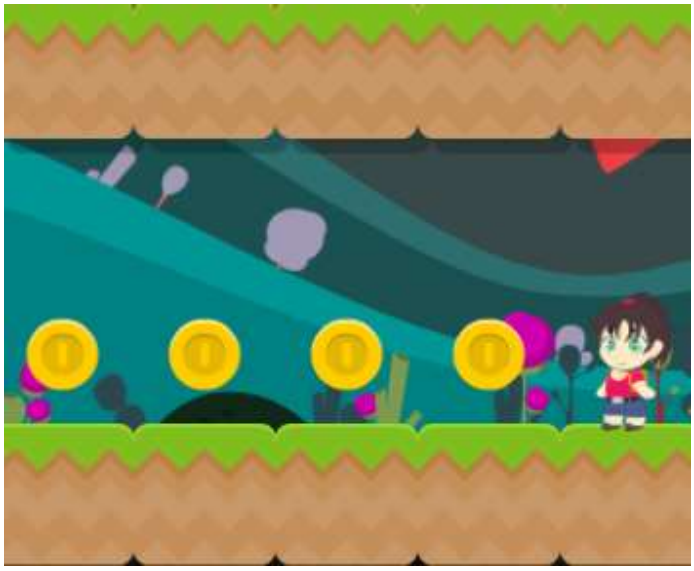

Figure 19 – Coins that can be picked up by the player in the game world

3.5.3 Puzzle System

A one-screen puzzle system was implemented as the final challenge in a given level in order to accommodate the metrics identified in Sections 3.4.2.1.3 and 3.4.2.2.1. As discussed in Section 3.4.2.1.3, it is hypothesized that players who score highly in the Cautiousness facet would spend more time contemplating a puzzle before attempting it. The puzzle game is played by moving coloured blocks onto matching coloured goal positions Figure 20. It is important to the game metric here that the puzzle is fully viewable before the player is even required to make the first move in order to allow them the time to deliberate.

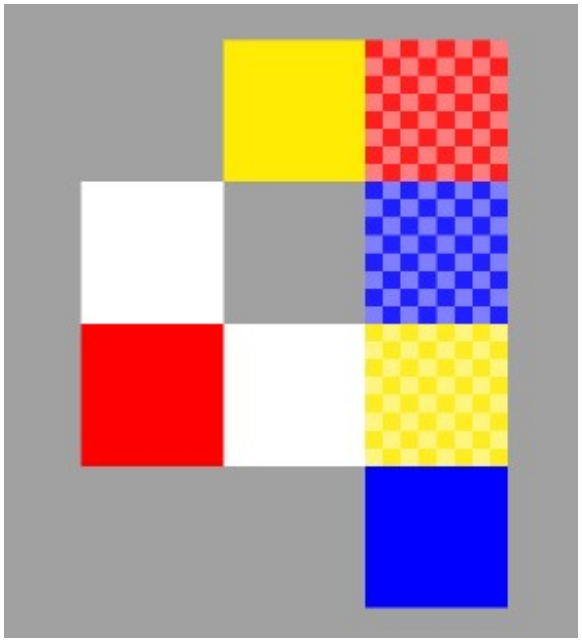

Figure 20 – Block pushing puzzle presented to the player after they make it to the level exit

As discussed in Section 3.4.2.2.1, it is hypothesized that players who score highly in the Intellect facet will choose more cerebrally challenging tasks. In order to accommodate this, the player is asked to choose the difficulty of the puzzle before they start it, where selecting a more difficult challenge increases the number of coloured blocks to be matched or expanding the board size (therefore increasing the total number of possible moves).

3.5.4 Leaderboard System

A leaderboard system was implemented in order to accommodate the metrics identified in Sections 3.4.2.1.6 and 3.4.2.3.4. As discussed in Section 3.4.2.3.4, it is hypothesized that players who score highly in the facet of Modesty will be less likely to share their score on a leaderboard. The leaderboard itself is populated with randomly generated entries so as to fake the social aspect of the leaderboard as seen in Figure 21. The player's score is highlighted in yellow while the randomly generated scores can never exceed the maximum amount of score that is able to be accumulated in the level.

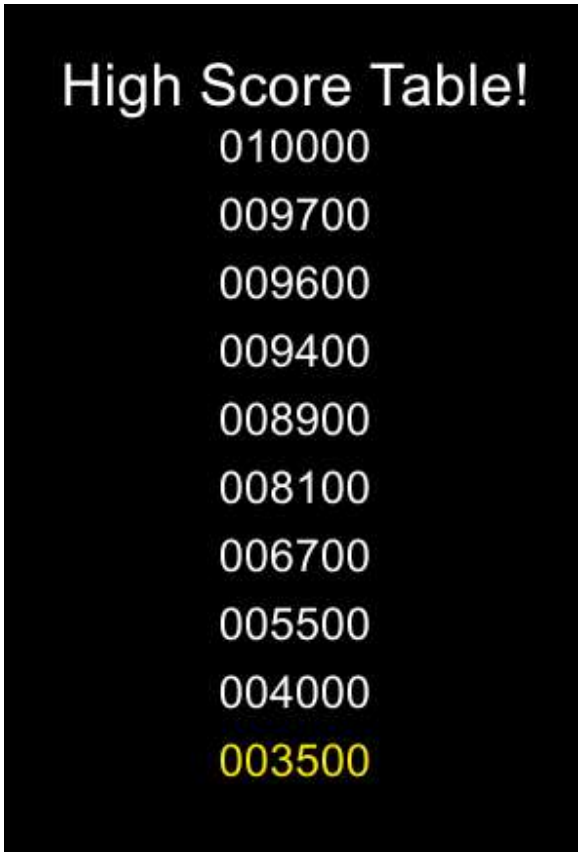

Figure 21 – The leaderboard filled with randomly generated numbers with the player's score highlighted in yellow

As discussed in Section 3.4.2.1.6, it is hypothesized that players who score highly in the facet of Dutifulness will avoid behaviour that is considered to be breaking the rules. The leaderboard system features a prompt that asks if the player would like to upload their score Figure 22. The player's score is displayed in a text box which can be edited by the player. Should they choose to edit this value, that false value will be displayed as their high score in the leaderboard.

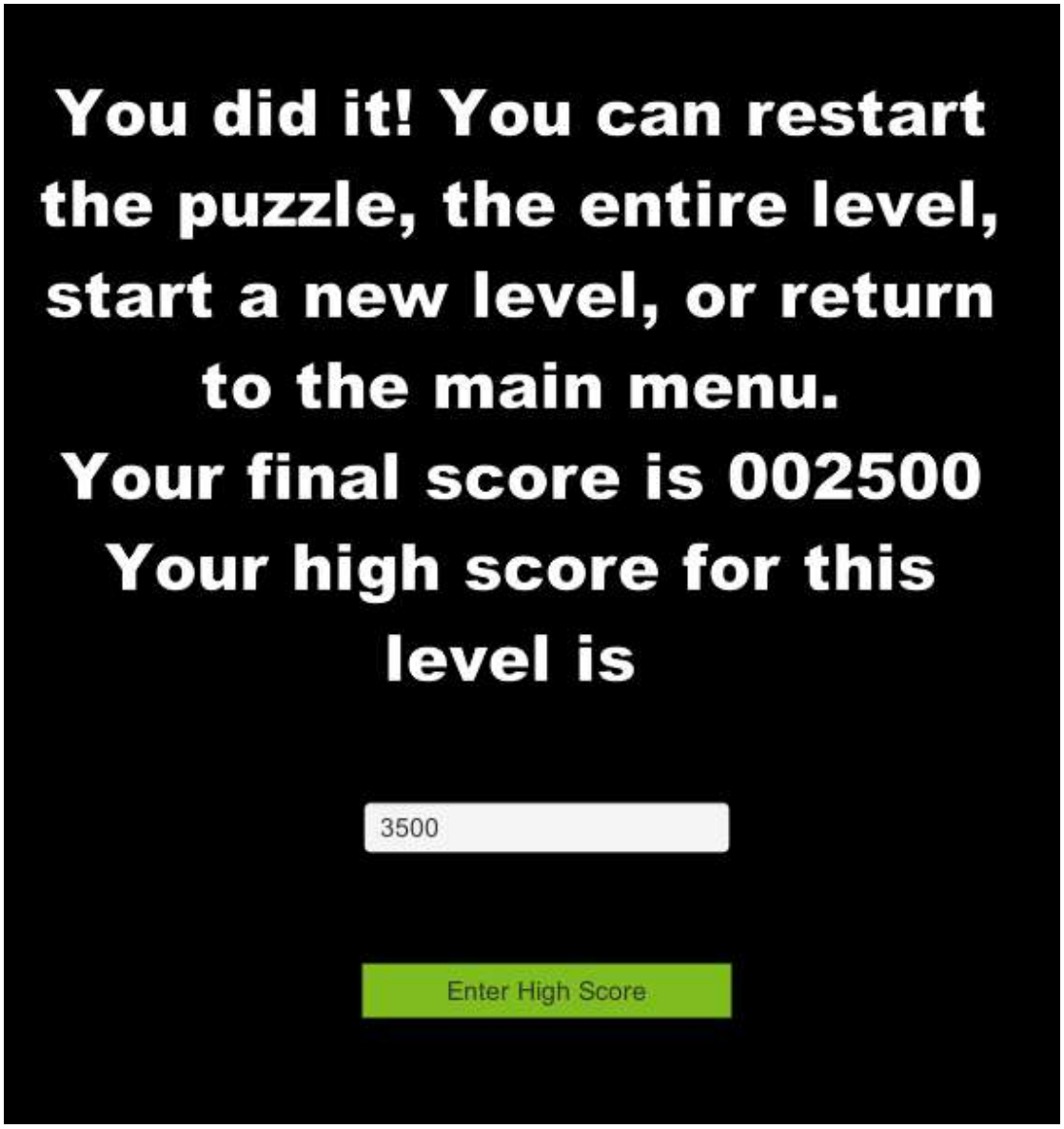

Figure 22 – Message given to the player informing them of their score along with editable text field displaying their score

3.5.5 Aesthetic System

A system that allows players to alter the look of the game as well as their player character was implemented in order to accommodate for the metrics described in Sections 3.4.2.2.2, 3.4.2.2.3, and 3.4.2.2.4. As discussed in Section 3.4.2.2.2, it is hypothesized that players who score highly in the Adventurousness facet will choose to replace their player character's looks, while Section 3.4.2.2.3 hypothesizes that players who score highly in the Artistic Interests facet will choose to

alter the looks of the game more often and Section 3.4.2.2.4 hypothesizes that players who score highly in the Imagination facet will prefer more fantastical looks for their character. The aesthetic system can be accessed from the main menu and presents the player with a menu as seen in Figure 23 which allows the player to select the visual theme of the levels they are playing as well as the skin of their player character in the form of a human, orc, alien, or dragon.

Figure 23 – Aesthetic selection screen where the player can select a theme for the level and their character

3.5.6 Between-player Gift System

A between-player gift system was implemented in order to accommodate the metrics described in Sections 3.4.2.3.1 and 3.4.2.3.2. A floating chest (Figure 24) will always be present near the player character at the beginning of the level. If the player interacts with the chest, they will be asked if they would like to accept a gift from another randomly selected player. This could take the form of a boon (bonus score) or a detriment (score subtraction) and is determined randomly (as a faked social response). As discussed in Section 3.4.2.3.1, it is hypothesized that players who score highly in the Trust facet are more likely to accept this gift.

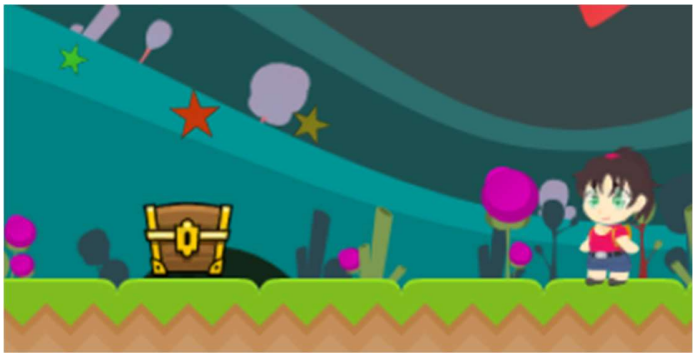

Figure 24 – Anonymous gift presented to the player near the beginning of the level

As discussed in Sections 3.4.2.3.2 hypothesized that players who score highly in the Altruism facet are more likely to give a good gift given a chance. Once the player reaches the exit of the level, they are asked if they would like to give a random person a gift (Figure 25). In reality, these buttons only collect data about the player's decision and does not actually cause any good or bad gifts to be sent out.

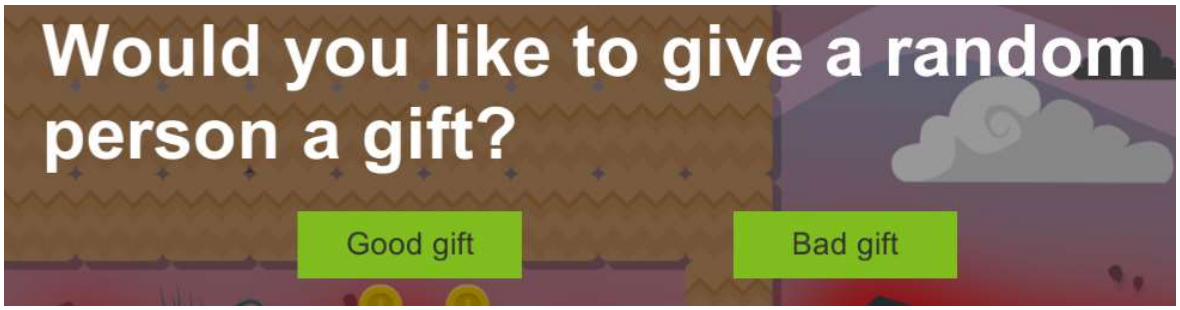

Figure 25 – Text prompt asking the player if they would like to pass on a gift

3.5.7 Non-playable Character (NPC) System

A non-playable character (NPC) system was implemented in order to accommodate the metrics described in Sections 3.4.2.4.1 and 3.4.2.4.2. The NPCs of the game are little hooded entities within the game world that do not move. When the player is within range of an NPC, they are given options to interact with them via one of four possible buttons (Figure 26). Clicking on a face button will display that face and cause the NPCs to mirror that face back at the player. As discussed in Section 3.4.2.4.1, it is hypothesized that players who score highly in the Cheerfulness facet will choose to utilize the happy face more often when interacting with NPCs.

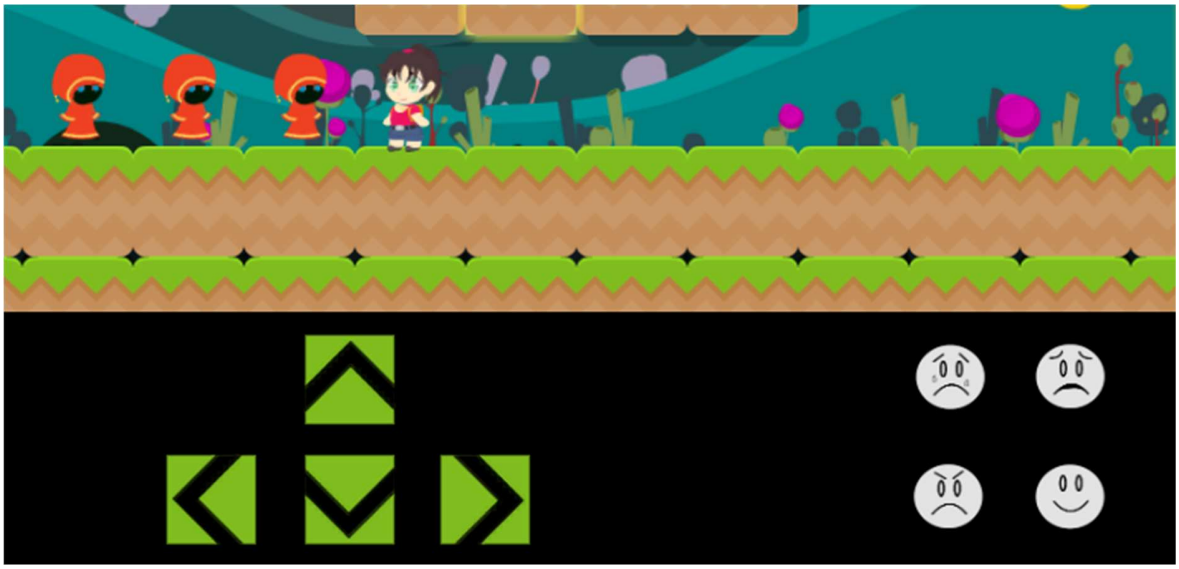

Figure 26 – Emote controls and NPCs that will react to them

As discussed in Section 3.4.2.4.2, it is hypothesized that players who score highly in the Gregariousness facet will choose to approach NPC groups more often than those who scored lowly. As part of this design, the NPCs are always designed to appear off of the necessary path to the exit while still being within view of the player such that if the player approaches the NPCs, they are making a conscious decision to go towards them.

3.5.8 Procedurally Generated Level System

A procedural generation system that creates random levels was implemented in the game in order to allow for increased replayability. This is not a criteria determined from the design process described in Section 3.4, but was deemed necessary to the video game. The traditional questionnaire tools are designed to consistently produce the same result (or as similar as a result as possible) over multiple testing sessions as a measure of its stability (Pittenger 2005; Burnett 2013). In order for the video game tool to achieve consistent testing results within an individual, the individual needs a fresh challenge every time they play which also helps keep them interested in the game (as a solved puzzle will be less engaging if replayed) (Despain et al. 2013; Jesse Schell 2008; Salmond 2016). Thus, the game world within the video game tool is procedurally generated from a series of hand designed sections of level. Each section is filled with several reusable items and is designed such that the player is always able to travel through it with the items that have

been provided. Players are able to restart a level and be presented with the same one they had seen before, but quitting out to the main menu will reset the random seed generation for that level. In doing so, the player can be presented with a fresh level layout every time they play the game but also be given the opportunity to replay a level if they felt like they had made an unfixable mistake or if they wanted to optimize their play in a level.

3.5.9 Random Number Generation

As discussed in Section 3.5.8 the procedural level generation system relies heavily on random number generation and was set up in such a way as to allow for players being able to replay a level if they wanted but also to offer them a new level if they left and returned to the game (or if they quit out to the main menu). In order to facilitate the same level being regenerated the random number generator was seeded (and the seed remembered). A small complication that arises from this is that certain aspects of the game need to be fully random even while playing the same level (such as the between-player gift system (Section 3.5.6)). In order to accommodate this, another unseeded random number generator was also used to allow for inconsistencies to exist between different instances of the same procedurally generated level.

As later discussed in Section 3.6.2.2 the seed for the randomly generated level is recorded and stored in an online repository. Using the exact same code base this would allow for a later reproduction of that level. The results of other truly random events were recorded separately so that an accurate recreation of a player's experience could be recreated if deemed necessary for any reason.

3.5.10 Summary of the Game's Implementation

Several systems were implemented based on the design described in Section 3.4 which forms the actual video game tool designed for personality assessment. This begins the process of answering SQ2 which is concerned with the opportunities and challenges of implementing a video game designed to heavily embed a theoretical model. The next section describes a deployment of the video game tool to obtain data that will help further refine the video game tool.

3.6 Deploying the Video Game Tool

The previous sections discussed the process of designing and implementing a video game from a theoretical groundwork. This section discusses how that design was tested. The next logical step is comparing the results from the designed video game as compared to contemporary personality data elicitation techniques. Due to limitations that will be discussed later, in Chapter 5, the designed video game does not have a metric that targets every single one of the 30 facets within the FFM. This begs the question of how the results compare at both a trait and facet level of personality.

3.6.1 Participant Selection

The video game and associated system was hosted in an online environment where anyone with an active internet connection could access it. Participants were recruited opportunistically by taking advantage of any avenue of contact available such as media releases and bulk mailing lists to students at the University of Tasmania, using social media such as Facebook and Twitter of those associated with the project, posters, flyers, and personal visits to lectures to advertise the study.

Two previous studies in this area had 24 and 44 participants each and still found significant correlations (Van Lankveld et al. 2009; Van Lankveld, Spronck, et al. 2011), and thus this project sought to obtain a minimum of 50 participants for this study to be able to draw significant conclusions.

Participants were not restricted in terms of age, but were required to confirm that they are over the age of 18 for legal reasons. Participants were asked to create a profile through a website which consisted of a username and a password. As part of the survey process, demographic information was also collected as follows: Participant age, gender, how long they have played video games, how often they player video games, and their highest education level.

3.6.2 Procedure

Participants were given an information sheet for review and a consent form to ascent to. These were presented digitally as they navigated to the website where the project was hosted. After

reading and agreeing to the consent form, participants were asked to provide a username and password.

The actual experimental procedure included two activities that were randomly ordered to account for order effect:

1. completion of the IPIP NEO-PI-R, and
2. completion of one (or more) gameplay sessions.

3.6.2.1 IPIP NEO-PI-R Questionnaire

Refer to the appendix for the full inventory. The questionnaire was delivered via a form developed on a third party website known as Typeform. Appended to the beginning of the questionnaire was a series of demographic questions that asked for the participant's age, gender, how long they have played video games, how often they player video games, and their highest education level (see Appendix 8.1).

3.6.2.2 Initial Gameplay Session

Participants were asked to play the game that has been designed. They were allowed to play for as much or as little time as they desired. Data was collected automatically within the game as the participant played the game and was stored in an online repository described below in Section 3.6.3. Each participant was given a different set of procedurally generated levels (with the seed for each level also being recorded in the online repository).

3.6.2.3 Gameplay/Questionnaire Order

The order of the questionnaire being administered and the game being played was alternated in order to account for the effects of priming as done by other studies in the area (Van Lankveld et al. 2009; Van Lankveld, Spronck, et al. 2011). This provided the benchmark for comparability to existing methods where the designed video game tool should at least match the results of the most dominant method of personality assessment if it is to be a useful tool.

3.6.3 System

The designed video game was hosted online on the NECTAR research cloud using access provided by the University of Tasmania. All of the data collected from the video game was also

stored on the same cloud instance in a MySQL database. The video game tool automatically and directly uploads recorded player data to the MySQL database in five second intervals.

3.6.4 Data Collection and Analysis

Data collection ran between 01/03/2017 and 19/06/2017 for a total of 108 days.

By the end of the data collection period, there were two sets of personality information:

1. Data from the video game
2. Data from the IPIP NEO-PI-R

Comparisons will be made between the groups in order to obtain information that will help answer SQ2. The data from the IPIP NEO-PI-R questionnaire survey will have to be scored and converted into a personality profile. The data from the video game will have to be processed into a form that is comparable to the results of the personality profile.

4. Results

The previous chapter discussed the research question and how it can be broken down into sub research questions. A review of the literature was used to identify a model of personality and subsequently a questionnaire as part of the answer to SQ1. A process was described to turn questionnaire items into game mechanics which in turn were conglomerated into a game design description which further helped to answer SQ1. The game design description was then implemented as a video game tool and serves as the first steps to answering SQ2. An evaluation process was described that would deploy the video game tool in a live scenario.

This chapter now seeks to discuss the results of that deployment in order to gain more insight into SQ1 and SQ2.

4.1 Demographics

In total, the number of respondents that both played the game and completed a personality survey was N = 43. The 43 respondents with complete data played a total of 139 levels of the game that yielded at least some analyzable data. Of the 43 respondents, 6 (13.95%) identified as female, 36 (83.72%) identified as male, and 1 (2.33%) identified as other gender. Figure 27 shows the distribution of respondents' experience with video games as self-reported during the questionnaire.

Figure 27 - Distribution of respondents' experience with games

Due to the nature of the tool used to conduct the survey for the IPIP NEO-PI-R questionnaire, it is impossible to determine how many potential respondents dropped out of the experiment at the point of completing the questionnaire. However, the video game tool recorded a total of 60 unique respondents, translating to 17 dropouts (28.33%) with 71.67% of respondents who played the video game tool also completing the questionnaire and providing usable data. Due to the small sample size, no data was trimmed for the purposes of outliers though some data from the video game tool was removed where there was inaction (for example not completing a level) and will be explicitly stated when it occurs in the following sections.

4.2 Questionnaire scoring

The data obtained from the questionnaire needs to be scored before it can be used as part of the analysis. Each item in the IPIP NEO-PI-R is answered with a Likert scale that rates how accurate that statement's relationship to the respondent is: "Very inaccurate", "Moderately accurate", "Neither accurate nor inaccurate", "Moderately accurate", and "Very accurate".

The official website for the International Personality Item Pool (IPIP) (Software 2014) describes a process to score the questionnaire. For +keyed items, the response "Very Inaccurate" is assigned a value of 1, "Moderately Inaccurate" a value of 2, "Neither Inaccurate nor Accurate" a 3,

"Moderately Accurate" a 4, and "Very Accurate" a value of 5. For –keyed items, the response "Very Inaccurate" is assigned a value of 5, "Moderately Inaccurate" a value of 4, "Neither Inaccurate nor Accurate" a 3, "Moderately Accurate" a 2, and "Very Accurate" a value of 1. The assigned values of a facet grouping can be added together to obtain a facet score which can then be added to other related facet scores to obtain a factor score.

For example, Participant 26's responses for the facet of Self-Discipline under the factor of Conscientiousness are shown in the Table 12 below. They answered "Neither Accurate nor Inaccurate" and "Moderately Accurate" for the first two items which are +keyed. This translates to an assigned value of 3 and 4 respectively. They also answered "Moderately Accurate" and "Neither Accurate nor Inaccurate" for the final two items which are –keyed. This translates to an assigned value of 2 and 3 respectively. Adding all of the assigned values together yields a score of 12 for the facet of Self-Discipline.

Facet	Questionnaire Item	Key	Participant ID 26's Response	Assigned Value
C5: Self-Discipline	Am always prepared.	+keyed	Neither Accurate nor Inaccurate	3
C5: Self-Discipline	Carry out my plans.	+keyed	Moderately Accurate	4
C5: Self-Discipline	Waste my time.	–keyed	Moderately Accurate	2
C5: Self-Discipline	Have difficulty starting tasks.	–keyed	Neither Accurate nor Inaccurate	3
Facet Score:				12

Table 12 – Example of Self-Discipline facet being scored

A full list of all facet scores calculated for all participants can be seen in Appendix. Further following the process described in (Software 2014), the facet scores can be added together to obtain factor scores. Table 13 shows a small sample of the factor scores calculated from the deployment data. The leftmost column is the anonymous ID given to the player, followed by columns for their Neuroticism, Extraversion, Openness, Agreeableness, and Conscientiousness scores. A full table with the personality factor scores of every respondent can be found in the Appendix.

Player ID	Neuroticism	Extraversion	Openness	Agreeableness	Conscientiousness
26	73	78	84	78	86
216	63	103	105	91	75
417	83	67	93	82	87
545	40	83	70	70	65
564	75	83	92	84	73
658	85	70	81	83	81
1016	99	67	71	84	58
1097	55	71	84	74	74
1219	67	71	71	68	80
1526	89	60	76	96	84
2186	77	72	97	95	85
2452	83	56	93	74	75
2675	75	66	90	99	84

Table 13 – Sample of factor scores from experiment participants

This section described the process of processing the data obtained from respondents into a series of factor scores that serve as the personality profile for that individual as well as the comparison point for the data derived from the video game tool. Figure 28 shows the distribution of personality Factor scores as calculated from the questionnaire results. The next section describes how video game data is broken down into a useable form for analysis.

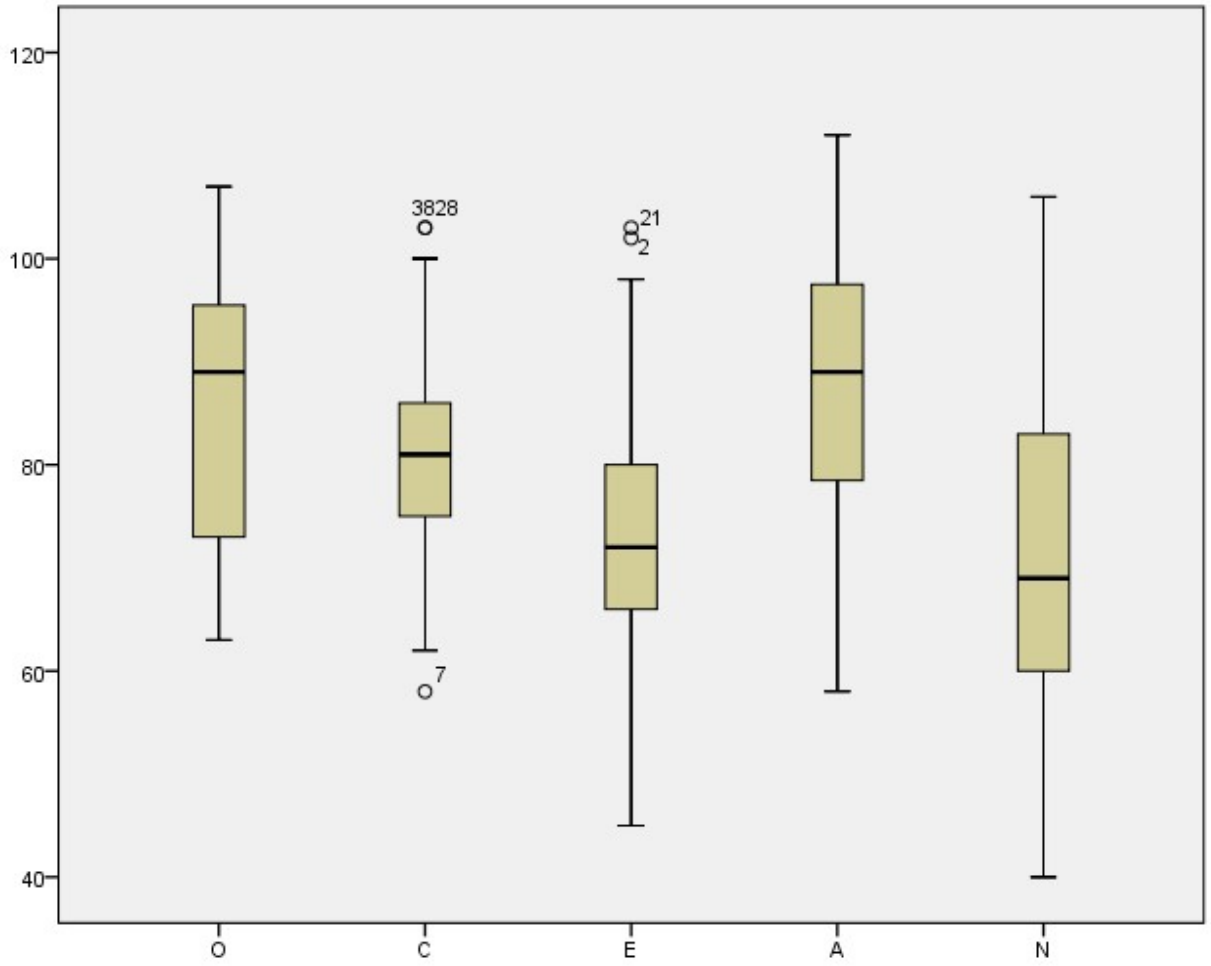

Figure 28 - Distribution of personality Factor scores

4.3 Game Data

The game metrics that were described in Section 3.4.2 were automatically collected as the player interacted with the video game tool. This section will describe the necessary processing to be conducted on the recorded play data in order to make meaningful comparisons with the personality scores calculated in the previous section. Pearson product-moment correlation coefficients are used to assess relationships between personality Factor scores derived from traditional questionnaires and the relevant game metrics for each facet. Two-tailed correlation was used in an attempt to minimize the risk of type 1 errors.

4.3.1 C2: Orderliness

The facet of Orderliness was tracked in-game using the number of inventory actions undertaken by the player per level (N = 133, M = 30.06, SD = 61.51) summarized in Figure 29. 6 levels were played with no interaction with the inventory system and were omitted.

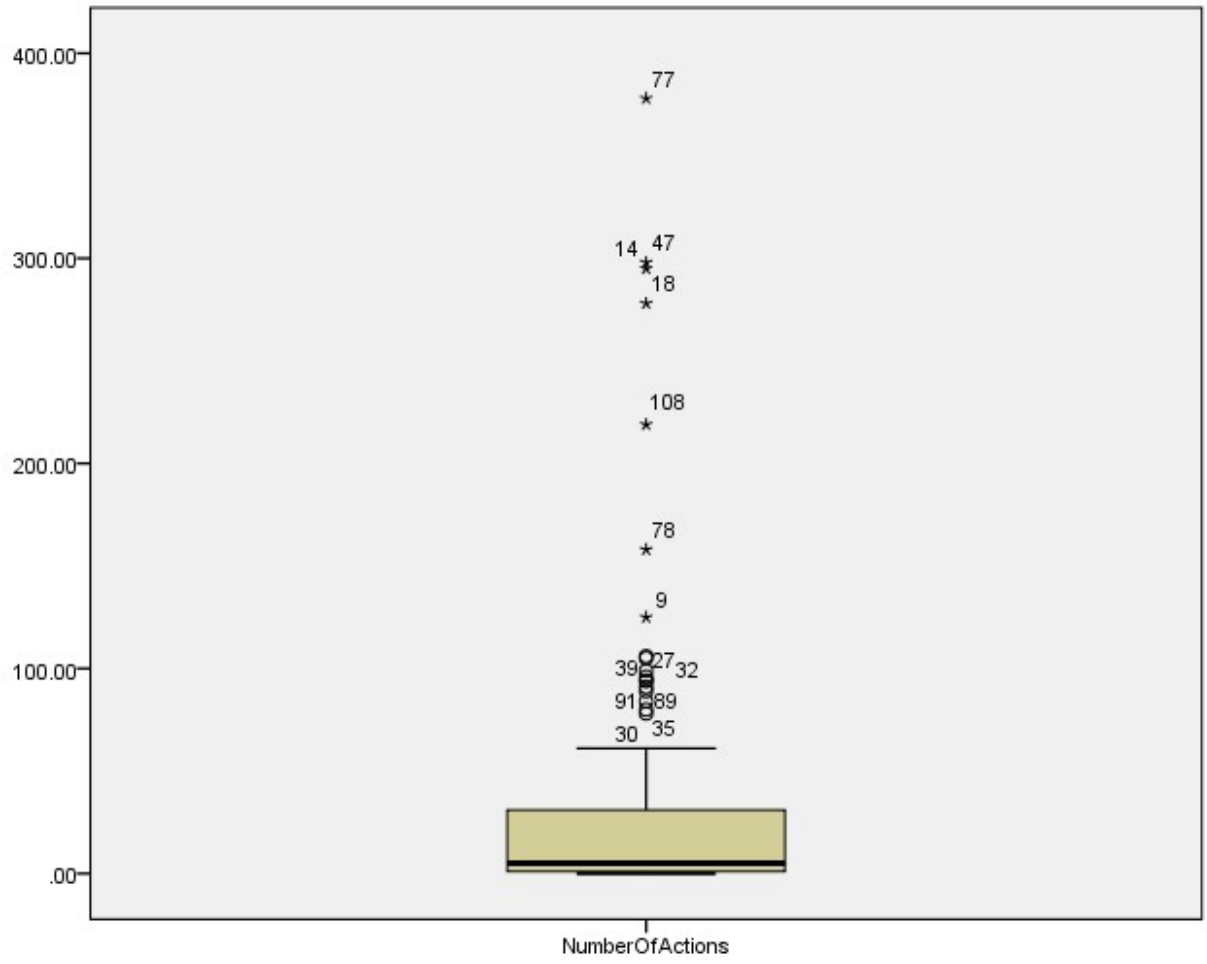

Figure 29 – Boxplot for number of inventory actions undertaken by the player per level

A Pearson product-moment correlation coefficient was computed to assess the relationship between personality Factor scores derived from traditional questionnaires and the average number of inventory actions undertaken by the player across all played levels.

Correlations

		O	C	E	A	N	Average no. Inventory Actions
AverageInvActs	Pearson Correlation	.083	-.219	-.013	-.009	-.084	1
	Sig. (2-tailed)	.595	.158	.933	.954	.593	
	N	43	43	43	43	43	43

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

There was no correlation between Openness and the number of inventory actions, $r(43) = .08$, $p > .05$. There was no correlation between Conscientiousness and the number of inventory actions, $r(43) = -.22$, $p > .05$. There was no correlation between Extraversion and the number of inventory actions, $r(43) = -.01$, $p > .05$. There was no correlation between Agreeableness and the number of inventory actions, $r(43) = -.01$, $p > .05$. There was no correlation between Neuroticism and the number of inventory actions, $r(43) = -.08$, $p > .05$.

Overall there was no significant correlation between any personality factor and the average number of inventory actions undertaken by the player. The game was unable to predict any factor of personality by recording the players inventory actions. This falls outside of the predictions for the project, as it was expected that this metric would form a correlation to Conscientiousness with no correlation to other factors.

4.3.2 C1: Self-Efficacy

The facet of Self-Efficacy was tracked in-game using the time spent in each completed level by the player (in seconds) (N = 49, M = 4364.12, SD = 3159.17) summarized in Figure 30. 90 levels were played but not completed and were omitted.

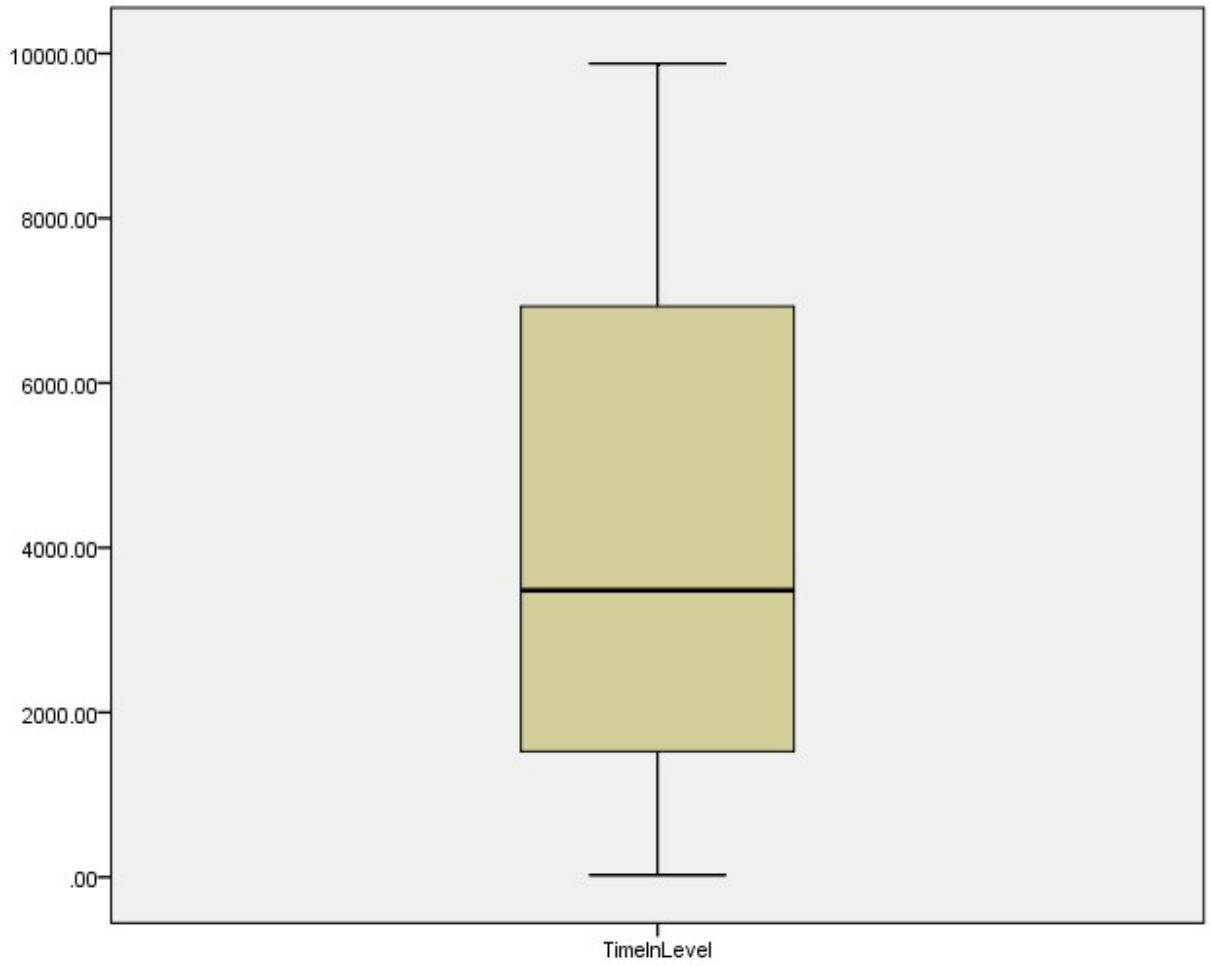

Figure 30 – Boxplot of time spent in each completed level by the player

A Pearson product-moment correlation coefficient was computed to assess the relationship between personality Factor scores derived from traditional questionnaires and the average time spent in a level by the player. 11 players did not complete a single level and were ignored for this calculation.

Correlations

		O	C	E	A	N	Average Time in Level
TimeInLvl	Pearson Correlation	-.159	-.106	-.218	.170	.040	1
	Sig. (2-tailed)	.384	.564	.231	.352	.829	
	N	32	32	32	32	32	32

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

There was no correlation between Openness and the time spent in the level, $r(32) = -.16$, $p > .05$. There was no correlation between Conscientiousness and the time spent in the level, $r(32) = -.11$, $p > .05$. There was no correlation between Extraversion and the time spent in the level, $r(32) = -.22$, $p > .05$. There was no correlation between Agreeableness and the time spent in the level, $r(32) = .17$, $p > .05$. There was no correlation between Neuroticism and the time spent in the level, $r(32) = .04$, $p > .05$.

Overall there was no significant correlation between any personality factor and the average time spent in a level by the player. The game was unable to predict any factor of personality by recording the amount of time a player invested in a level. This falls outside of the predictions for the project, as it was expected that this metric would form a correlation to Conscientiousness with no correlation to other factors.

4.3.3 C6: Cautiousness

The facet of Cautiousness was tracked in-game using the average moves made per second by the player in the puzzle component of the game (N = 25, M = .77, SD = .33) summarized in Figure 31. 114 levels were played without completing the puzzle and were omitted.

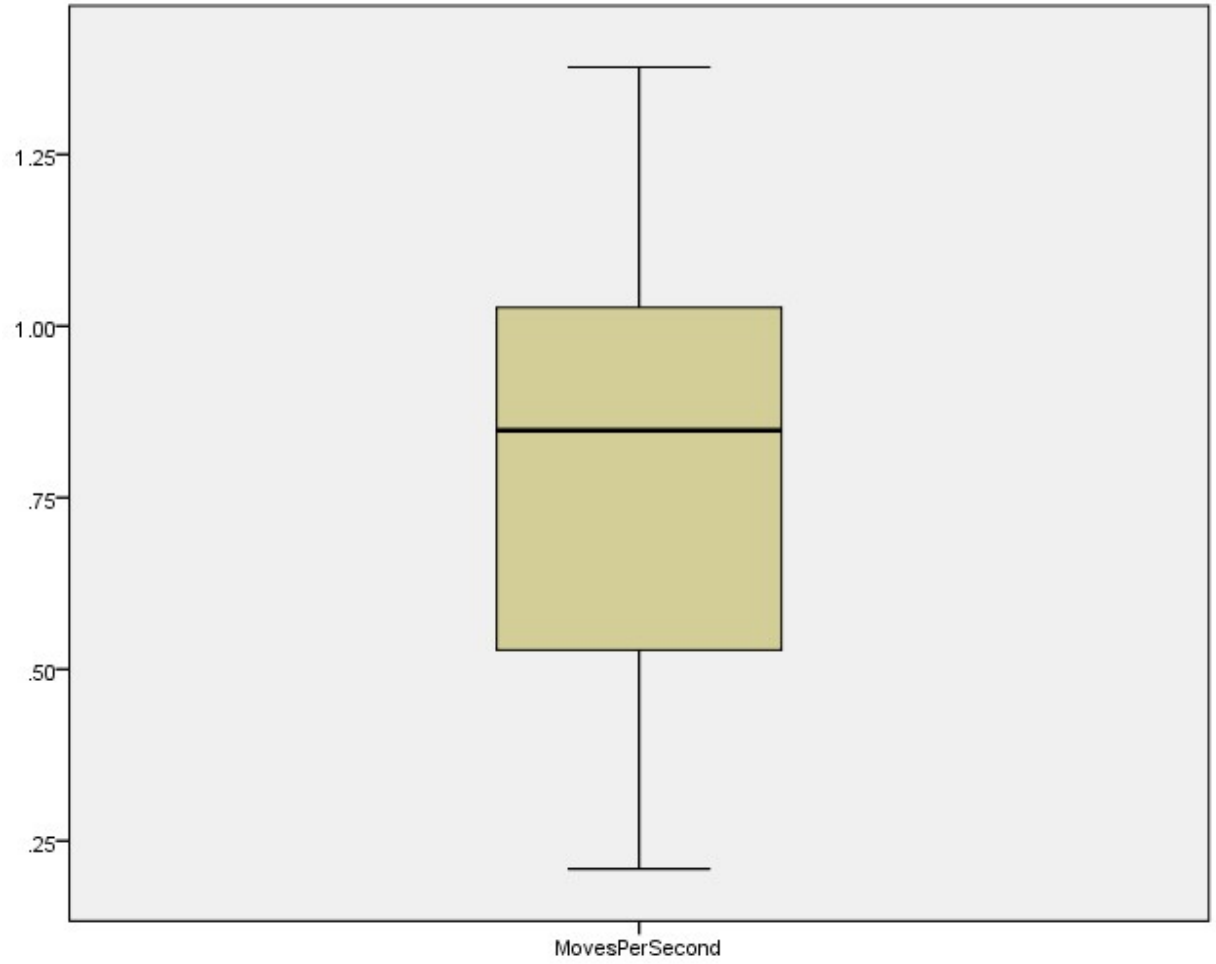

Figure 31 – Boxplot of average moves made per second by the player in puzzle

A Pearson product-moment correlation coefficient was computed to assess the relationship between personality Factor scores derived from traditional questionnaires and the average moves made by the player per second in the puzzle portion of the game across all levels. 24 players did not complete even a single instance of a puzzle and were ignored for this calculation.

Correlations

		O	C	E	A	N	Average Moves per Second
AverageMovePerTime	Pearson Correlation	-.083	.060	.145	.090	.253	1
	Sig. (2-tailed)	.736	.808	.554	.714	.297	
	N	19	19	19	19	19	19

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

There was no correlation between Openness and the moves made per second, $r(19) = -.08$, $p > .05$. There was no correlation between Conscientiousness and the moves made per second, $r(19) = -.06$, $p > .05$. There was no correlation between Extraversion and the moves made per second, $r(19) = .15$, $p > .05$. There was no correlation between Agreeableness and the moves made per second, $r(19) = .09$, $p > .05$. There was no correlation between Neuroticism and the moves made per second, $r(19) = .25$, $p > .05$.

Overall there was no significant correlation between any personality factor and the average number of moves made by the player during the puzzle portion of the game. The game was unable to predict any factor of personality by recording the players moves during the puzzle. This falls outside of the predictions for the project, as it was expected that this metric would form a correlation to Conscientiousness with no correlation to other factors.

4.3.4 C4: Achievement Striving

The facet of Achievement Striving was tracked in-game using the average number of coins collected by the player (compared to the maximum amount available in the level) per level (N = 139, M = .15, SD = .19) summarized in Figure 32.

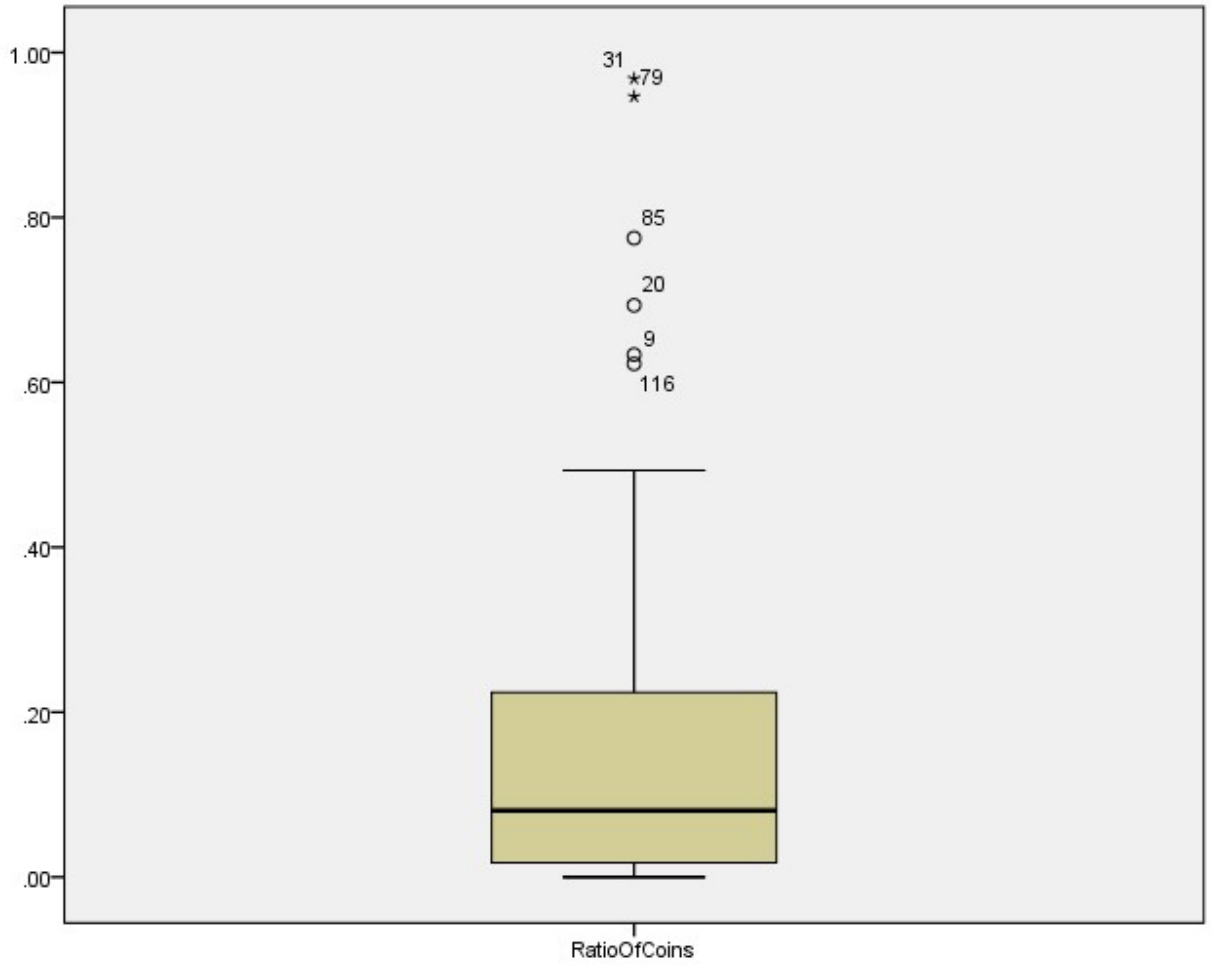

Figure 32 – Boxplot of number of coins collected by the player per level

A Pearson product-moment correlation coefficient was computed to assess the relationship between personality Factor scores derived from traditional questionnaires and the average number of coins collected by the player (compared to the maximum amount available in the level).

Correlations

		O	C	E	A	N	Average Ratio of Coins Collected
AverageCoins	Pearson Correlation	.120	-.178	.083	.079	-.142	1
	Sig. (2-tailed)	.443	.253	.596	.616	.363	
	N	43	43	43	43	43	43

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

There was no correlation between Openness and the number of coins collected, $r(43) = .12$, $p > .05$. There was no correlation between Conscientiousness and the number of coins collected, $r(43) = -.18$, $p > .05$. There was no correlation between Extraversion and the number of coins collected, $r(43) = .08$, $p > .05$. There was no correlation between Agreeableness and the number of coins collected, $r(43) = .08$, $p > .05$. There was no correlation between Neuroticism and the number of coins collected, $r(43) = -.14$, $p > .05$.

Overall there was no significant correlation between any personality factor and the average number of coins that a player collected (with respect to the maximum amount of coins available in that level). The game was unable to predict any factor of personality by recording the player's coin collecting behaviour. This falls outside of the predictions for the project, as it was expected that this metric would form a correlation to Conscientiousness with no correlation to other factors.

4.3.5 C5: Self-Discipline

The facet of Self-Discipline was tracked in-game using the number of items turned to pickups in each level (N = 139, M = 8.09, SD = 14.74) per level summarized in Figure 33.

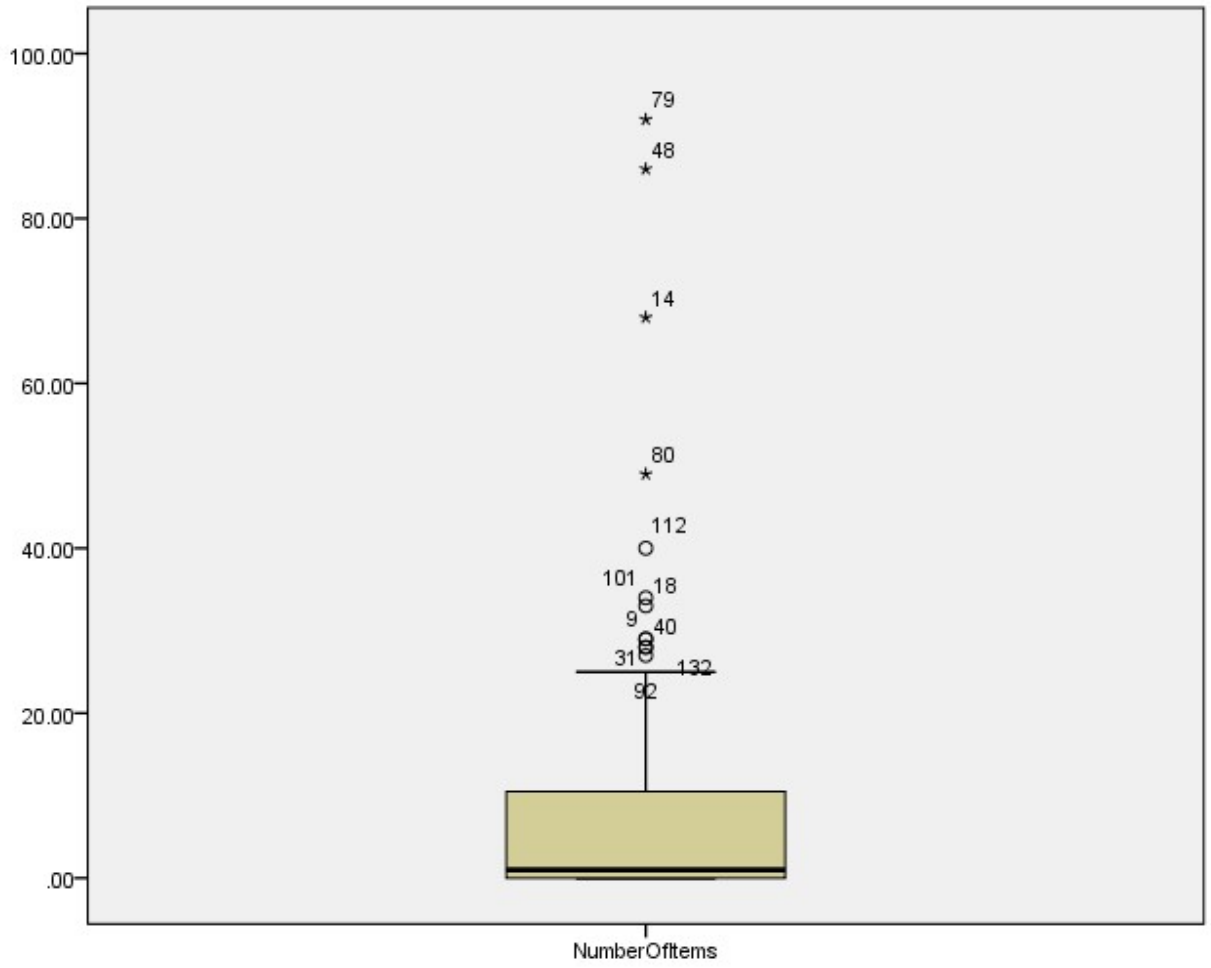

Figure 33 – Boxplot of the number of items turned to pickups per level

A Pearson product-moment correlation coefficient was computed to assess the relationship between personality Factor scores derived from traditional questionnaires and the average number of times the player turned items in the world into pickups.

Correlations

		O	C	E	A	N	Average no. of Items Turned to Pickups
AverageItemToPickup	Pearson Correlation	.116	-.213	.027	.015	-.082	1
	Sig. (2-tailed)	.458	.171	.862	.925	.603	
	N	43	43	43	43	43	43

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

There was no correlation between Openness and the average number of items turned to pickups, $r(43) = .12$, $p > .05$. There was no correlation between Conscientiousness and the average number of items turned to pickups, $r(43) = -.21$, $p > .05$. There was no correlation between Extraversion and the average number of items turned to pickups, $r(43) = .03$, $p > .05$. There was no correlation between Agreeableness and the average number of items turned to pickups, $r(43) = .02$, $p > .05$. There was no correlation between Neuroticism and the average number of items turned to pickups, $r(43) = -.08$, $p > .05$.

Overall there was no significant correlation between any personality factor and the average number of items turned to pickups by the player. The game was unable to predict any factor of personality by recording the player's interaction with items in the game world. This falls outside of the predictions for the project, as it was expected that this metric would form a correlation to Conscientiousness with no correlation to other factors.

4.3.6 C3: Dutifulness

The facet of Dutifulness was tracked in-game using the number of times the player cheated with respect to the amount of opportunities they had (N = 15, M = .13, SD = .35) summarized in Figure 34. 124 levels were played without seeing the screen where they would be allowed to cheat.

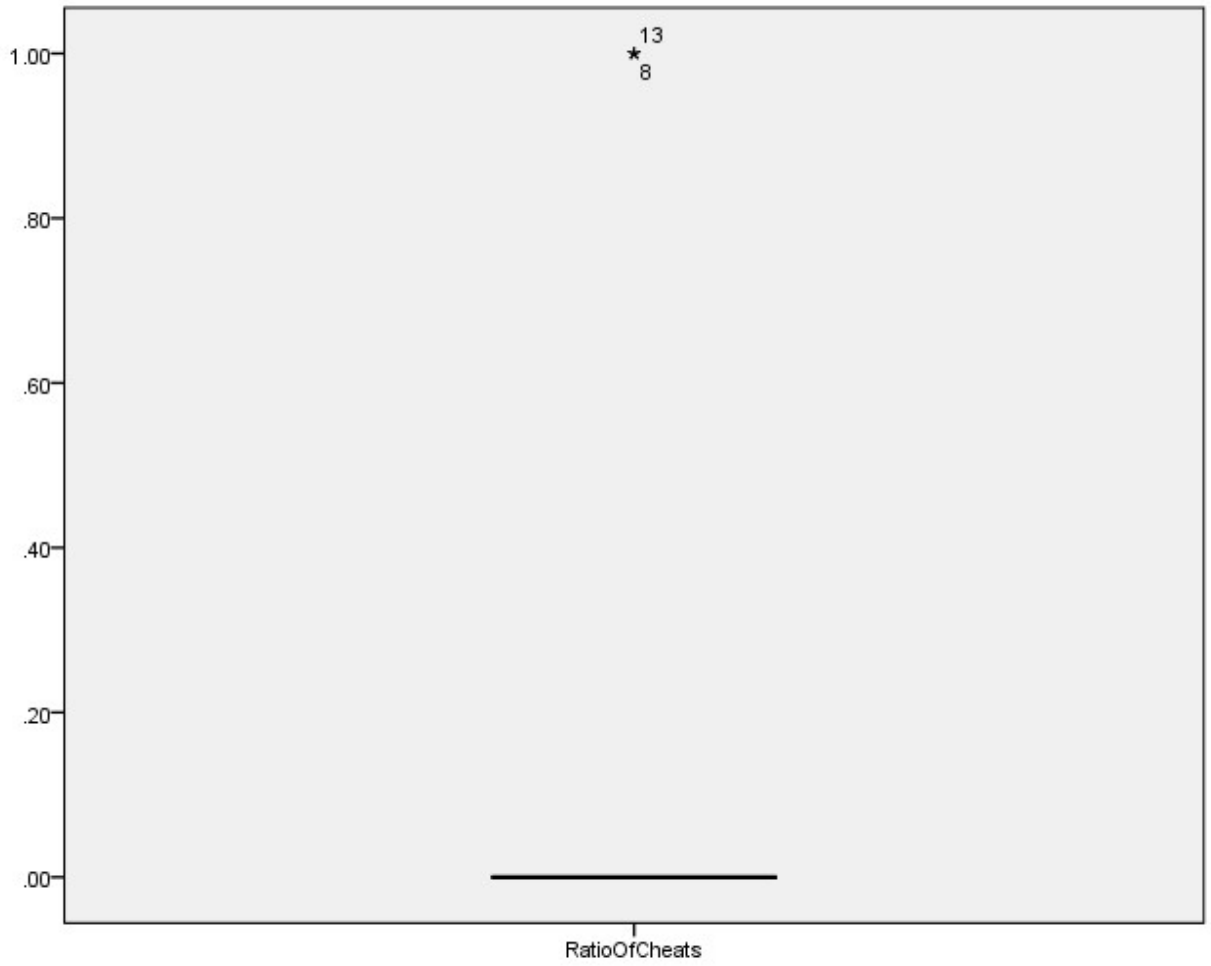

Figure 34 – Boxplot of the number of times the player cheated

A Pearson product-moment correlation coefficient was computed to assess the relationship between personality Factor scores derived from traditional questionnaires and the number of times the player cheated with respect to the amount of opportunities they had. 28 players never saw the screen that would have allowed the cheat to happen and were ignored for this calculation.

Correlations

		O	C	E	A	N	Cheat Ratio
CheatRatio	Pearson Correlation	-.155	-.197	-.462	-.164	-.088	1
	Sig. (2-tailed)	.580	.481	.083	.560	.756	
	N	15	15	15	15	15	15

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

There was no correlation between Openness and the chances to cheat, $r(15) = -.16$, $p > .05$. There was no correlation between Conscientiousness and the chances to cheat, $r(15) = -.20$, $p > .05$. There was no correlation between Extraversion and the chances to cheat, $r(15) = -.46$, $p > .05$. There was no correlation between Agreeableness and the chances to cheat, $r(15) = -.16$, $p > .05$. There was no correlation between Neuroticism and the chances to cheat, $r(15) = -.09$, $p > .05$.

Overall there was no significant correlation between any personality factor and the times the player cheated. The game was unable to predict any factor of personality by presenting the player with an opportunity to cheat and recording their actions. This falls outside of the predictions for the project, as it was expected that this metric would form a correlation to Conscientiousness with no correlation to other factors.

4.3.7 O5: Intellect

The facet of Dutifulness was tracked in-game using the average difficulty chosen by the player for the puzzle portion of the game (N = 49, M = 2, SD = .76) summarized in Figure 35. 90 levels were played without selecting a difficulty for the puzzle portion of the game.

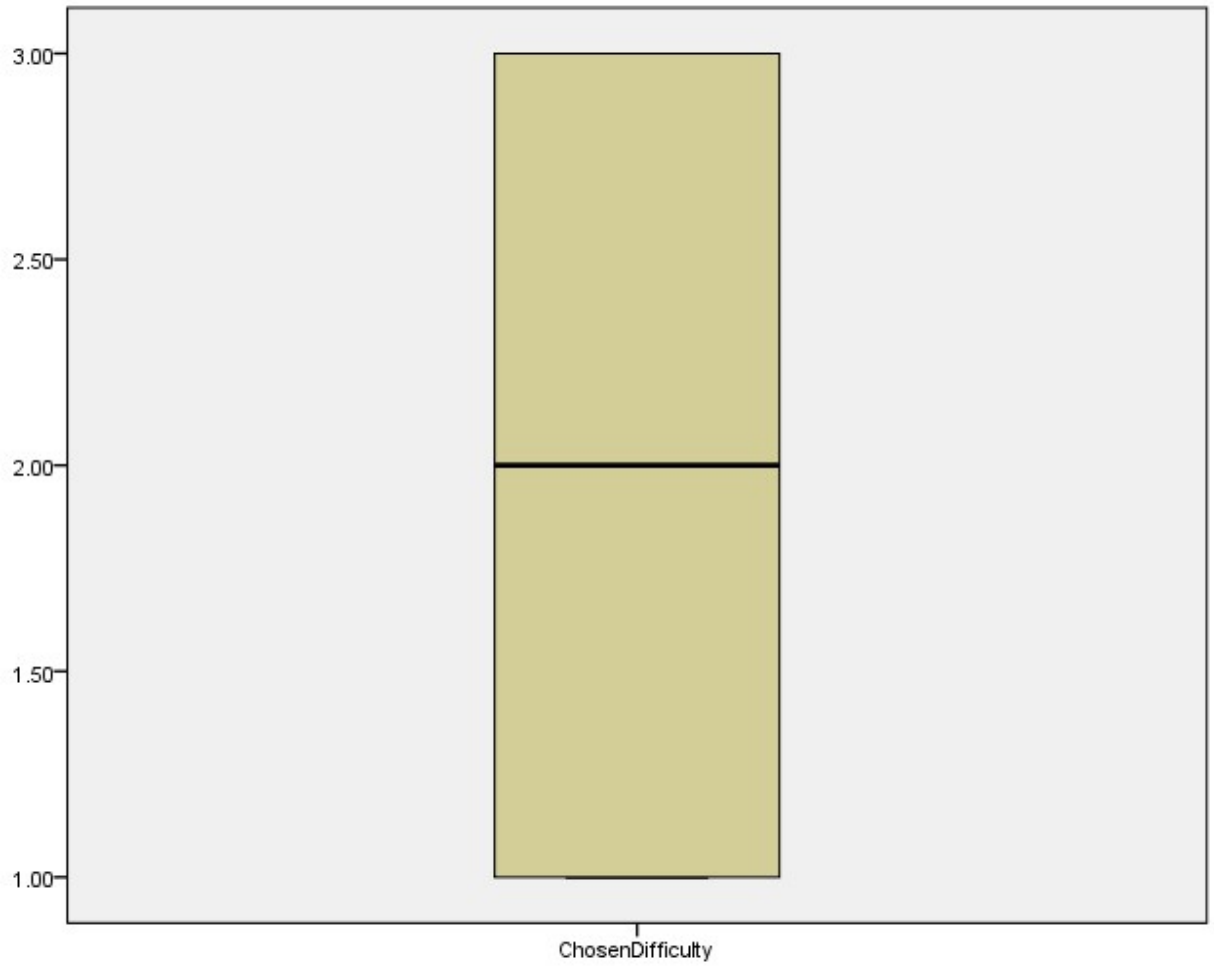

Figure 35 – Boxplot of the average puzzle difficulty chosen by the player

A Pearson product-moment correlation coefficient was computed to assess the relationship between personality Factor scores derived from traditional questionnaires and the average difficulty chosen by the player for the puzzle portion of the game. 11 players never selected a difficulty for the puzzle and were ignored for this calculation.

Correlations

		O	C	E	A	N	Average Difficulty of Puzzle
AveragePusherDiff	Pearson Correlation	.122	.163	-.318	-.109	.228	1
	Sig. (2-tailed)	.507	.374	.077	.551	.210	
	N	32	32	32	32	32	32

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

There was no correlation between Openness and the average difficulty chosen, $r(32) = .12$, $p > .05$. There was no correlation between Conscientiousness and the average difficulty chosen, $r(32) = .16$, $p > .05$. There was no correlation between Extraversion and the average difficulty chosen, $r(32) = -.32$, $p > .05$. There was no correlation between Agreeableness and the average difficulty chosen, $r(32) = -.11$, $p > .05$. There was no correlation between Neuroticism and the average difficulty chosen, $r(32) = .23$, $p > .05$.

Overall there was no significant correlation between any personality factor and the average difficulty chosen by the player for the puzzle portion of the game. The game was unable to predict any factor of personality by recording the player's preference for puzzle difficulty. This falls outside of the predictions for the project, as it was expected that this metric would form a correlation to Openness with no correlation to other factors.

4.3.8 O4: Adventurousness

The facet of Dutifulness was tracked in-game using the number of times the player made aesthetic changes between levels (N = 5, M = 2.60, SD = 1.34) summarized in Figure 36. Only 5 instances where a player played a level before making aesthetic changes for another level were recorded.

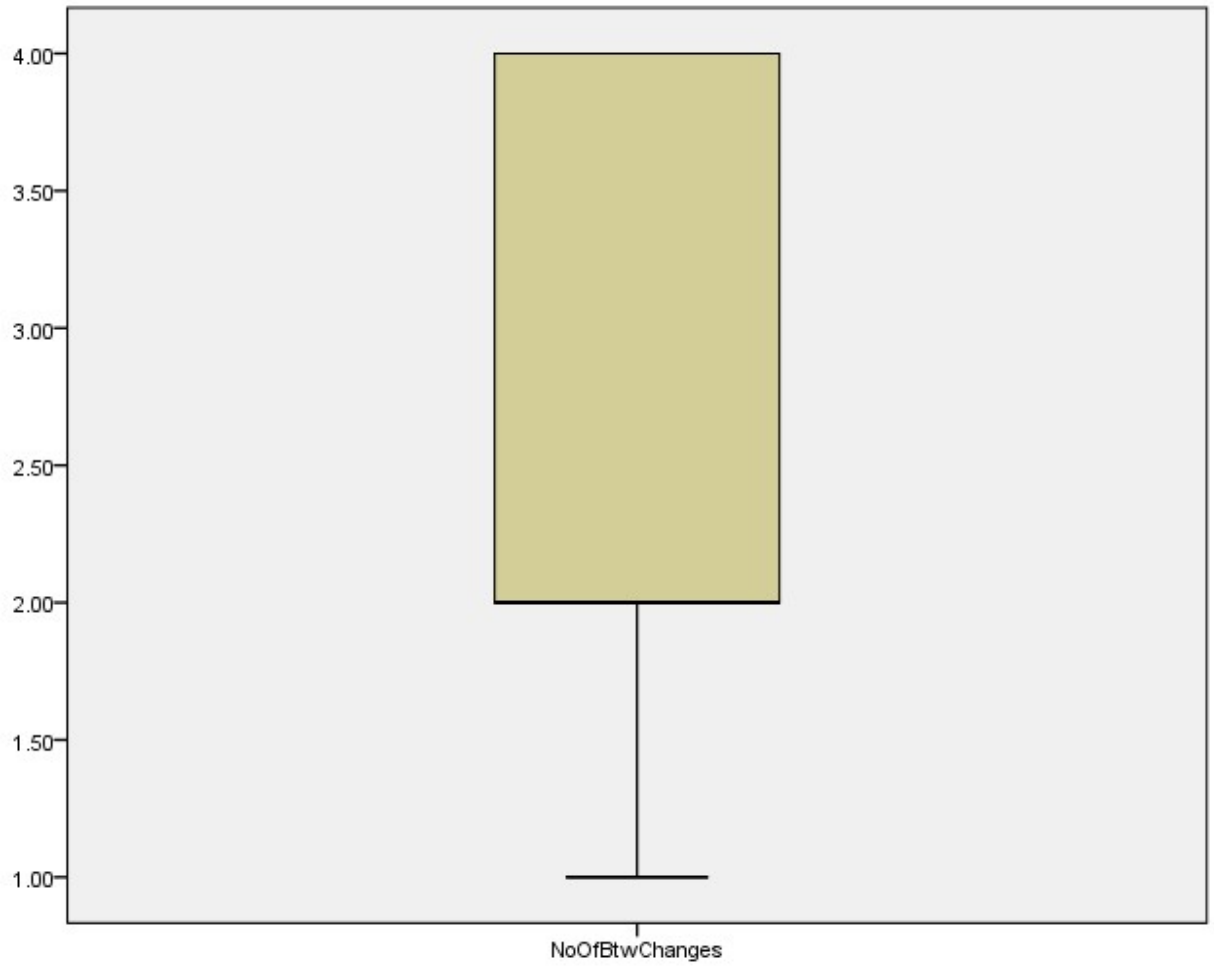

Figure 36 – Boxplot of the number of times the player made aesthetic changes between levels

A Pearson product-moment correlation coefficient was computed to assess the relationship between personality Factor scores derived from traditional questionnaires and the number of times the player made aesthetic changes between levels. Only 5 instances where a player played a level before making aesthetic changes for another level were recorded. All other players were ignored for this calculation.

Correlations

		O	C	E	A	N	Between Level Changes
BetweenChanges	Pearson Correlation	-.546	-.811	.245	-.906*	-.598	1
	Sig. (2-tailed)	.341	.096	.691	.034	.287	
	N	5	5	5	5	5	5

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

There was no correlation between Openness and the number of aesthetic changes between levels, $r(5) = -.55$, $p > .05$. There was no correlation between Conscientiousness and the number of aesthetic changes between levels, $r(5) = -.81$, $p > .05$. There was no correlation between Extraversion and the number of aesthetic changes between levels, $r(5) = .25$, $p > .05$. There was a significant correlation between Agreeableness and the number of aesthetic changes between levels, $r(5) = -.91$, $p = .03$. There was no correlation between Neuroticism and the number of aesthetic changes between levels, $r(5) = -.60$, $p > .05$.

Overall there were no significant correlations between the factors of Openness, Conscientiousness, Extraversion, and Neuroticism with the number of aesthetic changes made by the player between levels. There was a strong negative relationship between the factor of Agreeableness and the number of aesthetic changes made by the player. This falls outside of the predictions for the project, as it was expected that this metric would form a correlation to Openness with no correlation to other factors.

4.3.9 O2: Artistic Interests

The facet of Artistic Interests was tracked in-game using the number of times the player made aesthetic changes before starting the game for the first time (N = 14, M = 2.57, SD = 2.06) summarized in Figure 37. Only 14 players made aesthetic changes before starting the game for the first time.

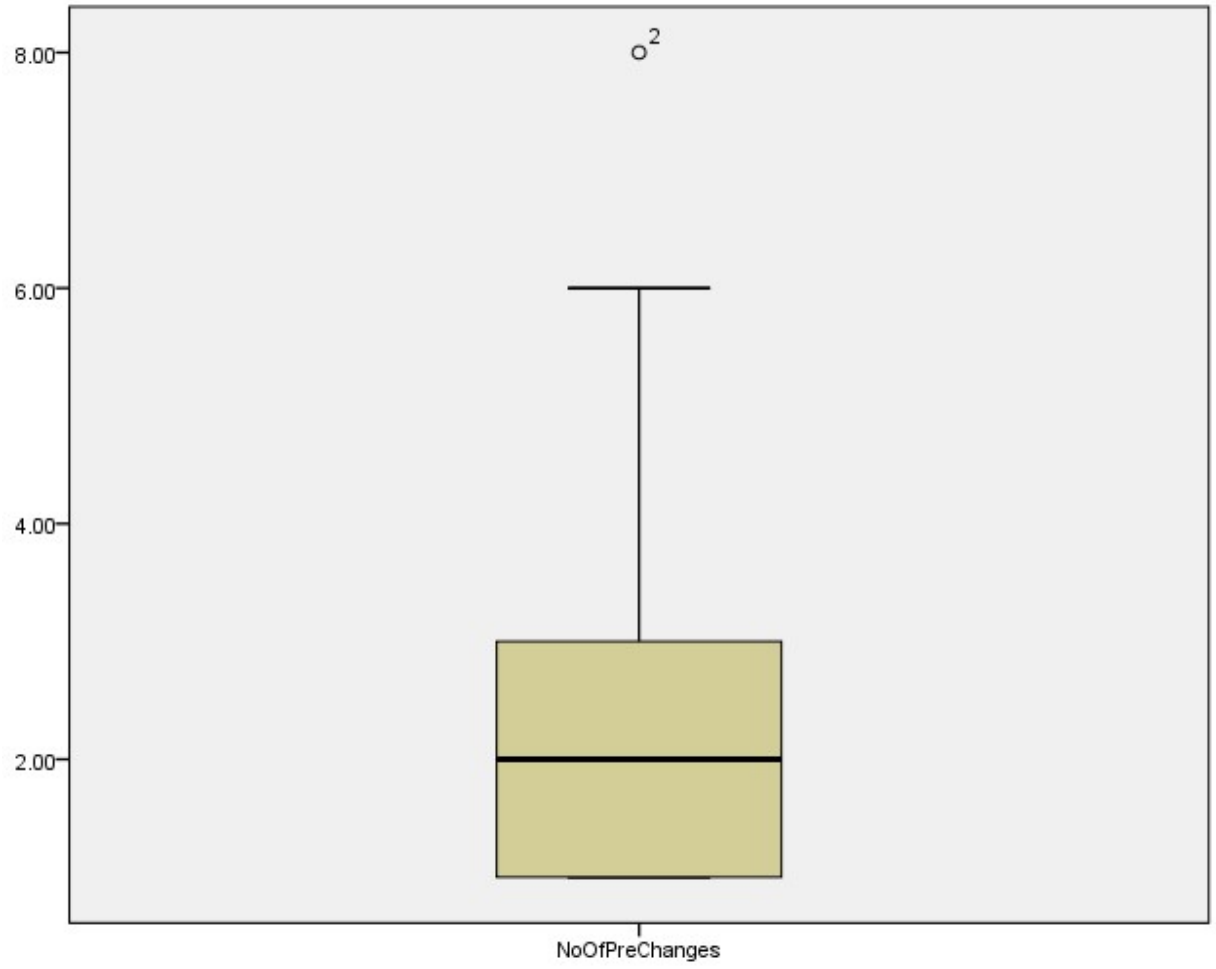

Figure 37 – Boxplot of the number of times the player made aesthetic changes before starting the game

A Pearson product-moment correlation coefficient was computed to assess the relationship between personality Factor scores derived from traditional questionnaires and the number of times the player made aesthetic changes before starting the game for the first time. 29 players did not make aesthetic changes before starting the game and were ignored for this calculation.

Correlations

		O	C	E	A	N	Pre-game Changes
PreChanges	Pearson Correlation	-.106	-.418	.026	.159	.321	1
	Sig. (2-tailed)	.719	.137	.930	.588	.262	
	N	14	14	14	14	14	14

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

There was no correlation between Openness and the number of aesthetic changes made before playing the game, $r(14) = -.11$, $p > .05$. There was no correlation between Conscientiousness and the number of aesthetic changes made before playing the game, $r(14) = -.42$, $p > .05$. There was no correlation between Extraversion and the number of aesthetic changes made before playing the game, $r(14) = .03$, $p > .05$. There was no correlation between Agreeableness and the number of aesthetic changes made before playing the game, $r(14) = .16$, $p > .05$. There was no correlation between Neuroticism and the number of aesthetic changes made before playing the game, $r(14) = .32$, $p > .05$.

Overall there was no significant correlation between any personality factor and the number of aesthetic changes made before the player started the game. The game was unable to predict any factor of personality through their selection of game aesthetic before the game began. This falls outside of the predictions for the project, as it was expected that this metric would form a correlation to Openness with no correlation to other factors.

4.3.10 O1: Imagination

The facet of Imagination was tracked in-game using the number of times the player made aesthetic choices preferring fantasy elements over mundane elements (N = 14, M = 1.86, SD = .86) summarized in Figure 38. 29 players did not make any aesthetic changes during their time with the game and were omitted.

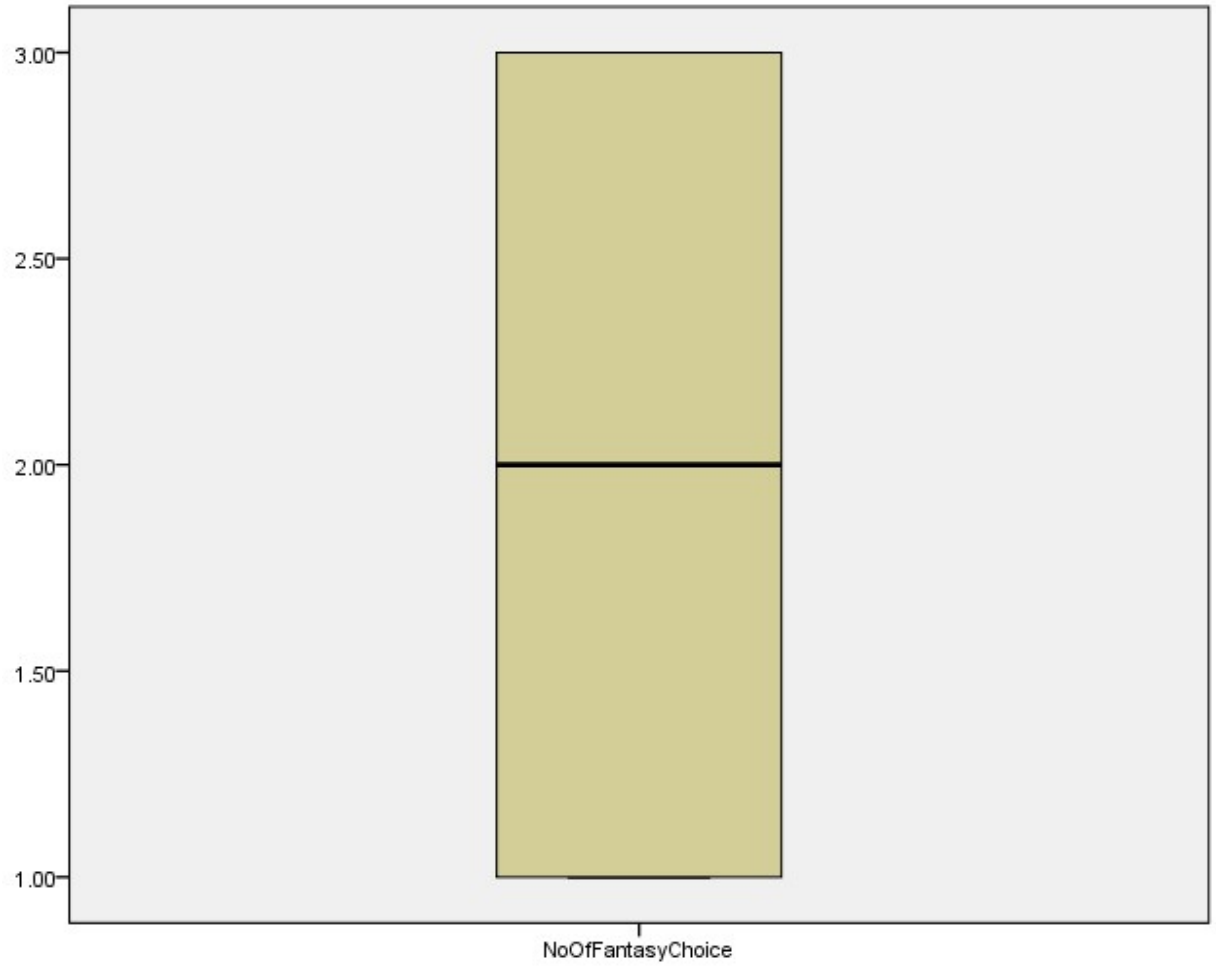

Figure 38 – Boxplot of the number of times the player preferred fantasy over mundane

A Pearson product-moment correlation coefficient was computed to assess the relationship between personality Factor scores derived from traditional questionnaires and the number of times the player made aesthetic choices preferring fantasy elements over mundane elements. 29 players did not make any aesthetic changes during their time with the game and were ignored for this calculation.

Correlations

		O	C	E	A	N	Choose Fantasy
Choose Fantasy	Pearson Correlation	-.561*	-.327	-.103	-.042	.095	1
	Sig. (2-tailed)	.037	.253	.725	.887	.746	
	N	14	14	14	14	14	14

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

There was a significant correlation between Openness and the number of fantasy choices made, $r(14) = -.57$, $p = .04$. There was no correlation between Conscientiousness and the number of fantasy choices made, $r(14) = -.33$, $p > .05$. There was no correlation between Extraversion and the number of fantasy choices made, $r(14) = -.10$, $p > .05$. There was no correlation between Agreeableness and the number of fantasy choices made, $r(14) = -.04$, $p > .05$. There was no correlation between Neuroticism and the number of fantasy choices made, $r(14) = .01$, $p > .05$.

Overall there were no significant correlations between the factors of Conscientiousness, Extraversion, Agreeableness, and Neuroticism with the number of times the player chose fantasy aesthetics over mundane aesthetics. There was a moderate negative relationship between the factor of Openness and the number of times the player preferred fantasy aesthetics over mundane aesthetics. This falls within the predictions for the project, expecting a relationship between this metric and Openness with no correlations to other factors.

4.3.11 A1: Trust

The facet of Trust was tracked in-game using the ratio of times the player accepted a random gift to the opportunities they received a random gift (N = 40, M = .89, SD = .22) summarized in Figure 39. 3 players never received a random gift and were omitted.

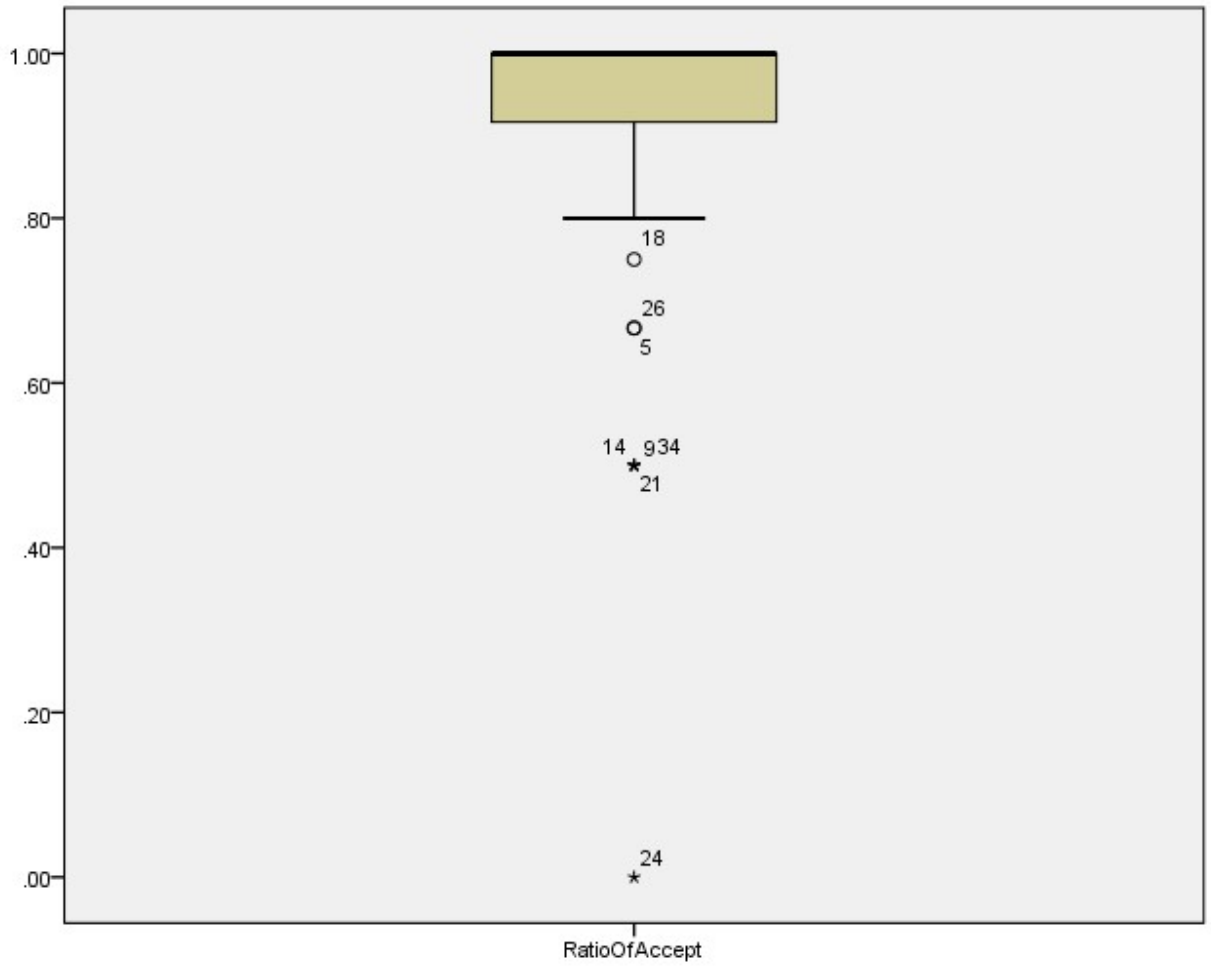

Figure 39 – Boxplot of the ratio of random gifts accepted to random gift opportunities

A Pearson product-moment correlation coefficient was computed to assess the relationship between personality Factor scores derived from traditional questionnaires and the number of times the player trusts and accepts the random gift. 3 players never received a random gift and were ignored for this calculation.

Correlations

		O	C	E	A	N	Chances To Trust
Chances To Trust	Pearson Correlation	.034	.178	-.065	.007	-.114	1
	Sig. (2-tailed)	.833	.271	.689	.965	.483	
	N	40	40	40	40	40	40

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

There was no correlation between Agreeableness and the number of times the player accepts the gift, $r(40) = .03$, $p > .05$. There was no correlation between Conscientiousness and the number of times the player accepts the gift, $r(40) = .18$, $p > .05$. There was no correlation between Extraversion and the number of times the player accepts the gift, $r(40) = -.07$, $p > .05$. There was no correlation between Agreeableness and the number of times the player accepts the gift, $r(40) = .01$, $p > .05$. There was no correlation between Neuroticism and the number of times the player accepts the gift, $r(40) = -.11$, $p > .05$.

Overall there was no significant correlation between any personality factor and the number of times the player accepted the random gift. The game was unable to predict any factor of personality through the player's choice in accepting gifts from other players. This falls outside of the predictions for the project, as it was expected that this metric would form a correlation to Agreeableness with no correlation to other factors.

4.3.12 A3: Altruism

The facet of Altruism was tracked in-game using the ratio of positive gifts given by the player to the opportunities they were presented (N = 28, M = .91, SD = .27) summarized in Figure 40. 15 players never got the opportunity to give a random gift and were omitted.

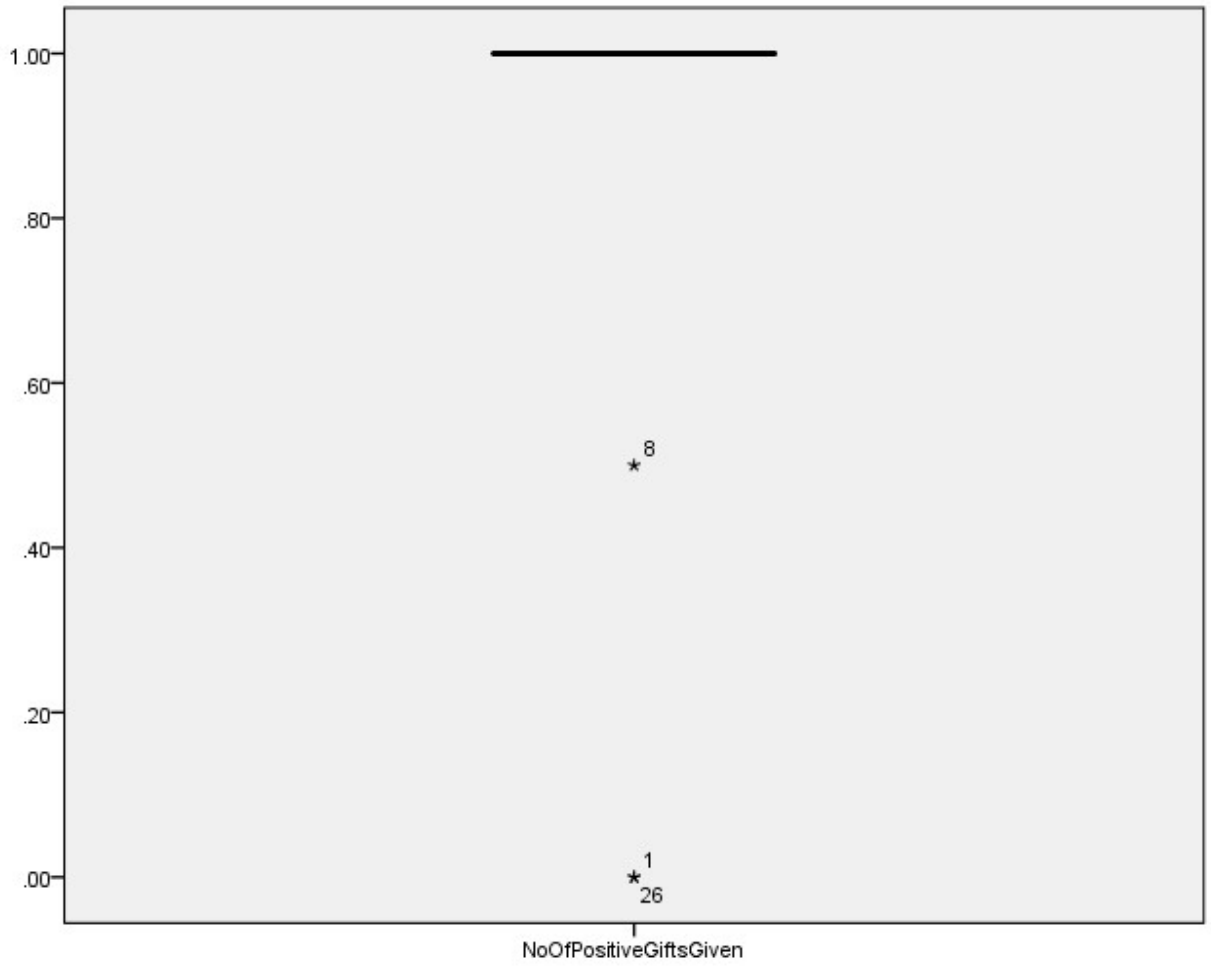

Figure 40 – Boxplot of the ratio positive gifts given by the player to opportunities they were presented

A Pearson product-moment correlation coefficient was computed to assess the relationship between personality Factor scores derived from traditional questionnaires and the ratio of positive gifts given by the player to the opportunities they were presented. 15 players never got the opportunity to give a random gift and were ignored for this calculation.

Correlations

		O	C	E	A	N	Chance To Give Good
ChanceGiveGood	Pearson Correlation	.077	-.009	.219	.393*	.035	1
	Sig. (2-tailed)	.697	.966	.262	.039	.859	
	N	28	28	28	28	28	28

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

There was no correlation between Agreeableness and the number of times the player gives positive gifts, $r(28) = .08$, $p > .05$. There was no correlation between Conscientiousness and the number of times the player positive good gifts, $r(28) = -.01$, $p > .05$. There was no correlation between Extraversion and the number of times the player gives positive gifts, $r(28) = .22$, $p > .05$. There was a significant correlation between Agreeableness and the number of times the player gives positive gifts, $r(28) = .39$, $p = .04$. There was no correlation between Neuroticism and the number of times the player gives positive gifts, $r(28) = .06$, $p > .05$.

Overall there were no significant correlations between the factors of Openness, Conscientiousness, Extraversion, and Neuroticism with the number of times the player chose to give positive gifts. There was a weak positive relationship between the factor of Agreeableness and the number of times the player gave good gifts. This falls within the predictions for the project, expecting a relationship between this metric and Openness with no correlations to other factors.

4.3.13 A4: Morality

The facet of Morality was tracked in-game using the ratio of negative gifts given by the player to the opportunities they were presented (N = 28, M = .09, SD = .27) summarized in Figure 41. 15 players never got the opportunity to give a random gift and were omitted.

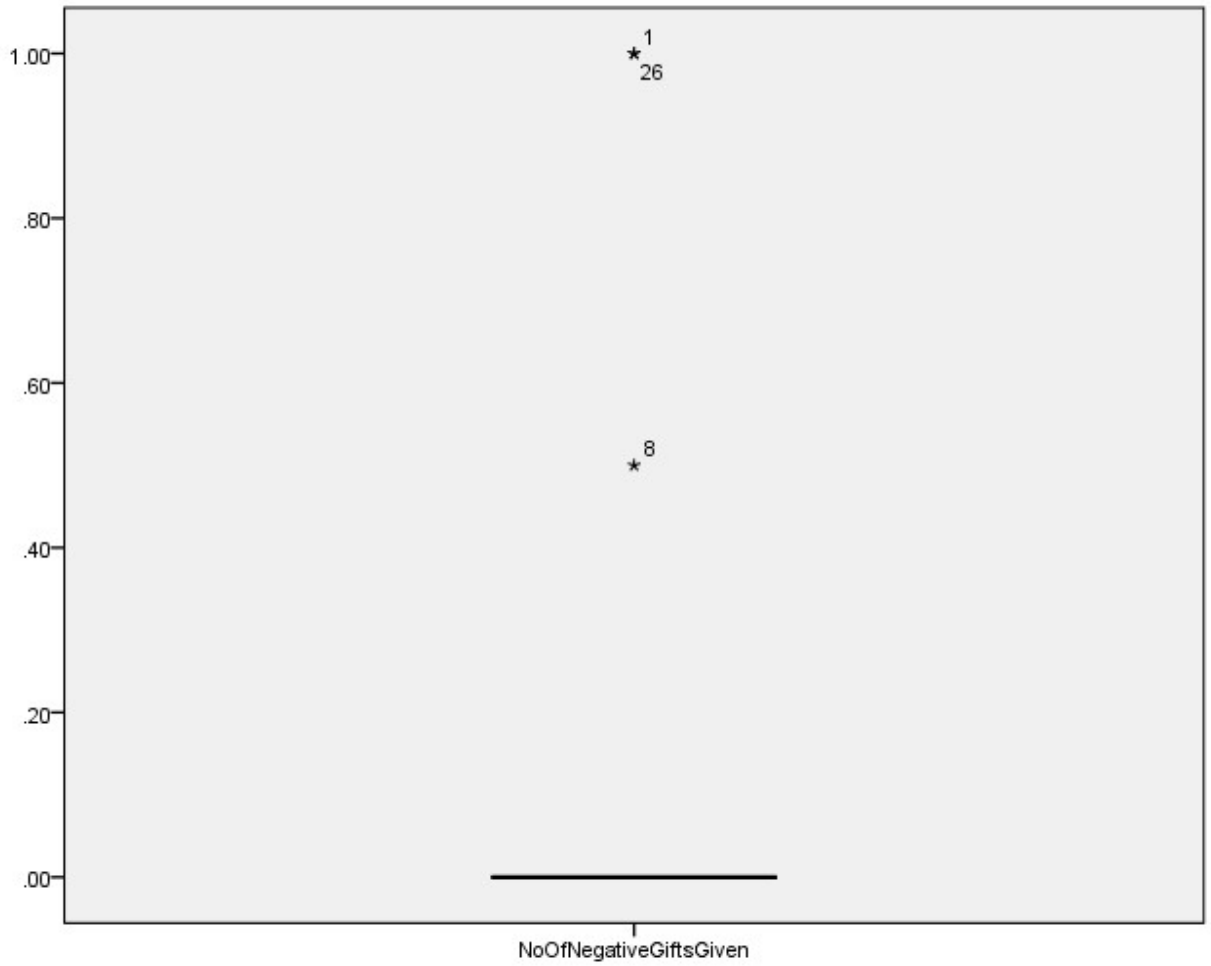

Figure 41 – Boxplot of the ratio negative gifts given by the player to the opportunities they were presented

A Pearson product-moment correlation coefficient was computed to assess the relationship between personality Factor scores derived from traditional questionnaires and the ratio of negative gifts given by the player to the opportunities they were presented. 15 players never got the opportunity to give a random gift and were ignored for this calculation.

Correlations

		O	C	E	A	N	Chance To Give Bad
ChanceGiveBad	Pearson Correlation	-.077	.009	-.219	-.393*	-.035	1
	Sig. (2-tailed)	.697	.966	.262	.039	.859	
	N	28	28	28	28	28	28

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

There was no correlation between Openness and the number of times the player gives negative gifts, $r(28) = -.08$, $p > .05$. There was no correlation between Conscientiousness and the number of times the player gives negative gifts, $r(28) = .01$, $p > .05$. There was no correlation between Extraversion and the number of times the player gives negative gifts, $r(28) = -.22$, $p > .05$. There was a significant correlation between Agreeableness and the number of times the player gives negative gifts, $r(28) = -.39$, $p = .04$. There was no correlation between Neuroticism and the number of times the player gives negative gifts, $r(28) = -.06$, $p > .05$.

Overall there were no significant correlations between the factors of Openness, Conscientiousness, Extraversion, and Neuroticism with the number of times the player chose to give negative gifts. There was a weak negative relationship between the factor of Agreeableness and the number of times the player gave good gifts. This falls within the predictions for the project, expecting a relationship between this metric and Openness with no correlations to other factors.

4.3.14 A5: Modesty

The facet of Modesty was tracked in-game using the number of times the player chooses to upload their high score (N = 19, M = .68, SD = .54) summarized in Figure 42. 24 players did not complete the game to a point of being given this option and were omitted.

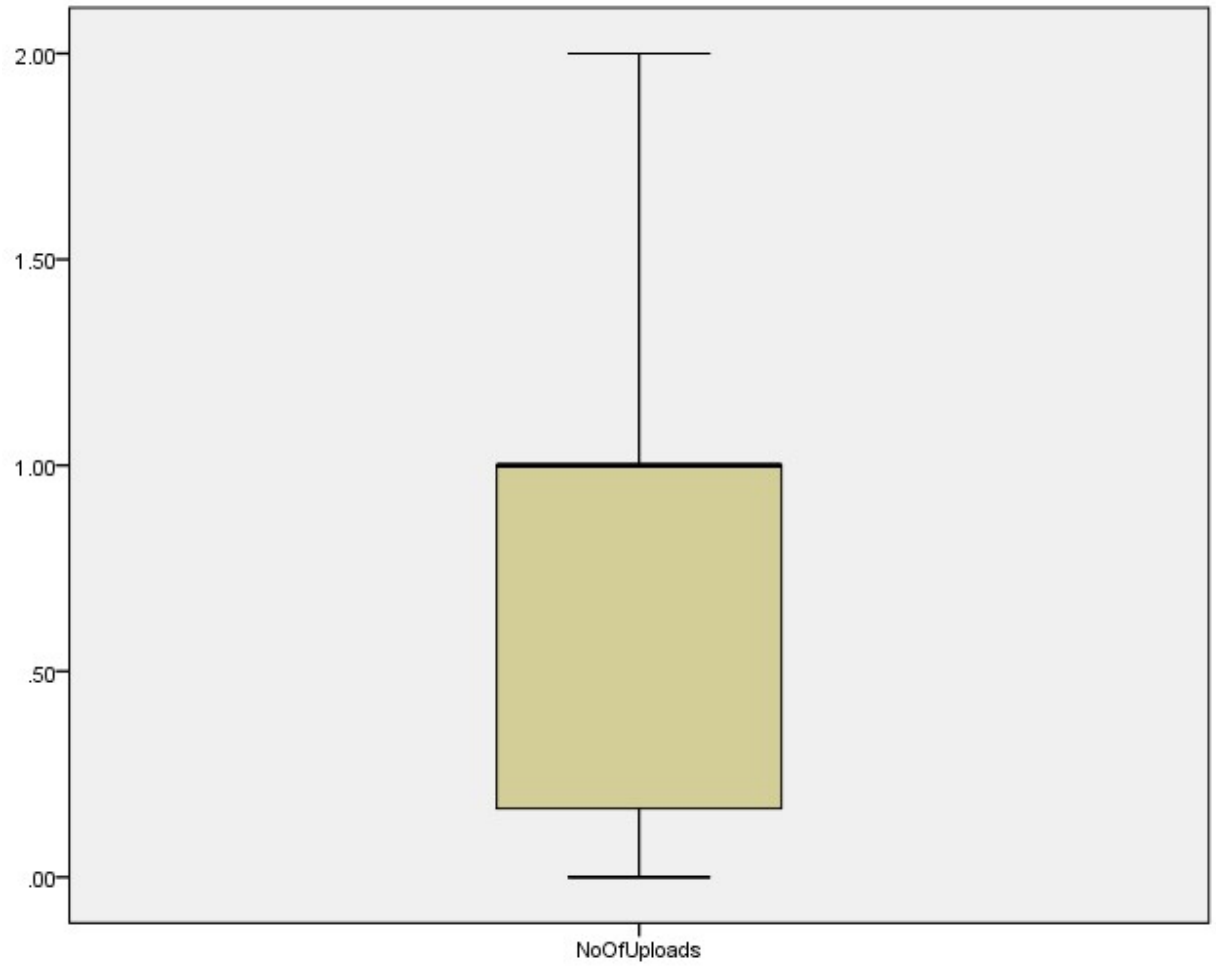

Figure 42 – Boxplot of the number of times the player uploaded their high score

A Pearson product-moment correlation coefficient was computed to assess the relationship between personality Factor scores derived from traditional questionnaires and the number of times the player chooses to upload their high score with respect to the chances they had. 24 players did not complete the game to a point of being given this option and ignored for this calculation.

Correlations

		O	C	E	A	N	Uploads
Uploads	Pearson Correlation	.189	-.052	.163	-.078	.026	1
	Sig. (2-tailed)	.437	.831	.504	.752	.914	
	N	19	19	19	19	19	19

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

There was no correlation between Openness and the number of times the player uploads their score, $r(19) = .19$, $p > .05$. There was no correlation between Conscientiousness and the number of times the player uploads their score, $r(19) = -.05$, $p > .05$. There was no correlation between Extraversion and the number of times the player uploads their score, $r(19) = .16$, $p > .05$. There was no correlation between Agreeableness and the number of times the player uploads their score, $r(19) = -.08$, $p > .05$. There was no correlation between Neuroticism and the number of times the player uploads their score, $r(19) = .03$, $p > .05$.

Overall there was no significant correlation between any personality factor and the number of times the player uploaded their high score with respect to the chances they had. The game was unable to predict any factor of personality by offering the opportunity to upload a high score to a leaderboard. This falls outside of the predictions for the project, as it was expected that this metric would form a correlation to Agreeableness with no correlation to other factors.

4.3.15 E6: Cheerfulness

The facet of Cheerfulness was tracked in-game using the ratio of times the player chose to utilize the Smile emotes to the number of times they used an emote (N = 13, M = .76, SD = 31) summarized in Figure 43. 30 players never used an emote and were omitted.

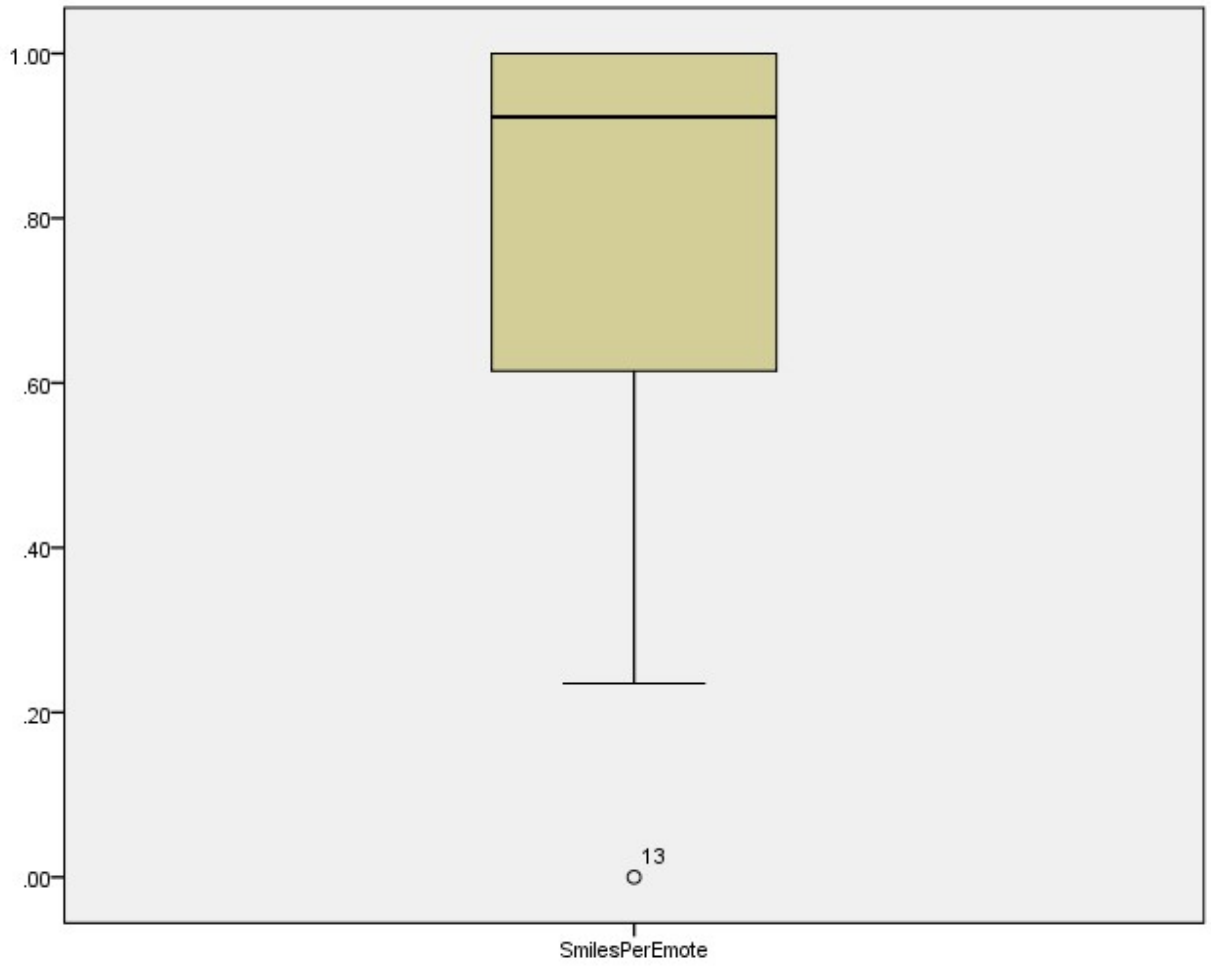

Figure 43 – Boxplot of the average number of times the player used Smile emote

A Pearson product-moment correlation coefficient was computed to assess the relationship between personality Factor scores derived from traditional questionnaires and the average number of times the player chose to utilize the Smile emotes over other available options. 30 players never used an emote and were ignored for this calculation.

Correlations

		O	C	E	A	N	Average no. of Smile Emotes
AverageSmileEmote	Pearson Correlation	.120	-.136	-.179	.131	.655*	1
	Sig. (2-tailed)	.696	.658	.559	.669	.015	
	N	13	13	13	13	13	13

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

There was no correlation between Openness and the average number of Smile emotes used, $r(13) = .12$, $p > .05$. There was no correlation between Conscientiousness and the average number of Smile emotes used, $r(13) = -.14$, $p > .05$. There was no correlation between Extraversion and the average number of Smile emotes used, $r(13) = -.18$, $p > .05$. There was no correlation between Agreeableness and the average number of Smile emotes used, $r(13) = .13$, $p > .05$. There was a significant correlation between Neuroticism and the average number of Smile emotes used, $r(13) = .66$, $p = .02$.

Overall there were no significant correlations between the factors of Openness, Conscientiousness, Extraversion, and Agreeableness with the average number of Smile emotes used by the player. There was a moderate positive relationship between the factor of Neuroticism and the average number of Smile emotes used by the player. This falls outside of the predictions for the project, as it was expected that this metric would form a correlation to Extraversion with no correlation to other factors.

4.3.16 E2: Gregariousness

The facet of Gregariousness was tracked in-game using the number of times players visited NPCs (N = 139, M = .50, SD = 1.12) per level summarized in Figure 44.

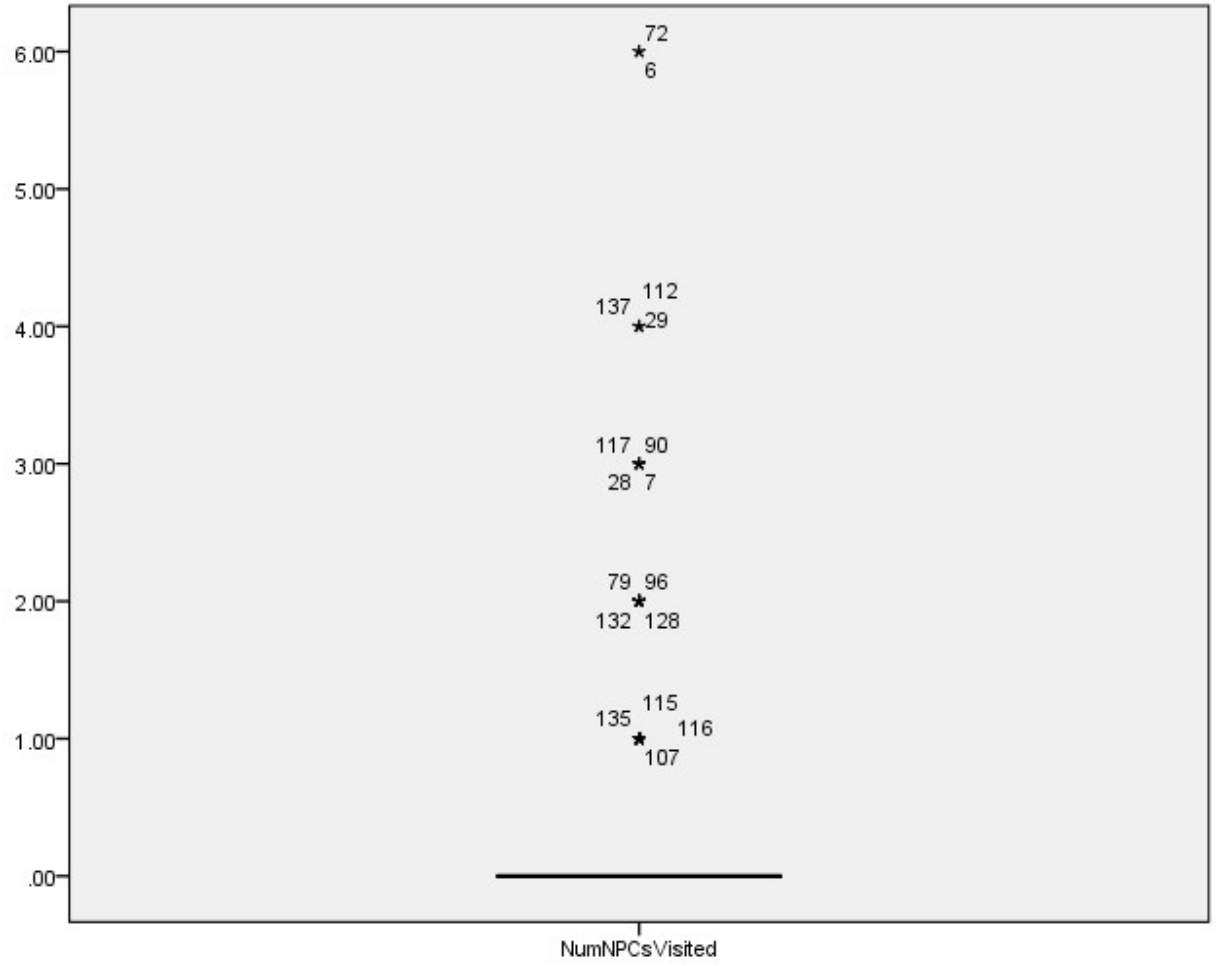

Figure 44 – Boxplot of the number of times the player visited NPCs

A Pearson product-moment correlation coefficient was computed to assess the relationship between personality Factor scores derived from traditional questionnaires and the number of times players visited NPCs with respect to the number of levels they played.

Correlations

		O	C	E	A	N	NPC Visits
NPC Visits	Pearson Correlation	.137	.228	-.121	.150	.035	1
	Sig. (2-tailed)	.381	.141	.441	.336	.823	
	N	43	43	43	43	43	43

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

There was no correlation between Openness and the number of times the player visited NPCs, $r(43) = .14$, $p > .05$. There was no correlation between Conscientiousness and the number of times the player visited NPCs, $r(43) = .23$, $p > .05$. There was no correlation between Extraversion and the number of times the player visited NPCs, $r(43) = -.12$, $p > .05$. There was no correlation between Agreeableness and the number of times the visited NPCs, $r(43) = .15$, $p > .05$. There was no correlation between Neuroticism and the number of times the player visited NPCs, $r(43) = .04$, $p > .05$.

Overall there was no significant correlation between any personality factor and the number of times the player visited NPCs per level played. The game was unable to predict any factor of personality by recording the player's reaction to NPCs. This falls outside of the predictions for the project, as it was expected that this metric would form a correlation to Extraversion with no correlation to other factors.

5. Discussion

The previous chapter presented the results of the experiment conducted to test the video game tool. This chapter will now reflect on the processes of this research in order to contextualize those results. The research question of the project was:

> What are the opportunities and challenges in creating a serious game to support the identification of player personality characteristics?

This research question is further broken down into the following sub-research questions:

1. What are the opportunities and challenges in *designing* a serious game to support the identification of player personality characteristics?
2. What are the opportunities and challenges in *implementing* a serious game to support the identification of player personality characteristics?

This discussion will begin in the design phase, discussing the methods used to derive a game design from a personality model in relation to SQ1 and then conclude in the implementation, discussing the systems that were implemented into the final video game tool in relation to SQ2. This discussion and the lessons learned from this process will answer both sub-research questions which in turn answers the research question and provide the starting point to improve the process for future work.

5.1 Reflection on the Design

This section will discuss all of the steps taken during the design process of the video game tool that begins from selecting a model of personality to be used as the basis for the video game up to the point of the final game design.

5.1.1 Personality Model Selection

The first step in the process of this project was the selection of a personality model which begins answering SQ1. The goal here was to select the model that best suited the needs of the research (see Section 3.1). To this end, a literature review was undertaken to highlight the best model of

personality for the project's goals of creating a game capable of making generalized measurements of an individual's personality (see Section 3.1).

As described in Section 3.1, The most prominent models of personality are the MBTI, MMPI, HEXACO, and FFM (Harkness et al. 2014; Myers et al. 1998; Lee & Ashton 2015; McCrae & Costa 2012). These models normally differentiate themselves around their applied fields, for example, the MBTI is largely used in corporations (Isachsen & Berens 1998), the MMPI in clinical and forensic settings (Pope et al. 2006), and the FFM and HEXACO model in academia (John & Srivastava, 1999). The MBTI suffers from severe criticisms to its validity and the MMPI's primary focus is in clinical psychology which removes them from consideration. The factored models of personality both model general personality, but the FFM is more prominently used over the HEXACO model as shown in the systematic literature review. For those reasons (summarized in Table 14), the FFM was chosen as the model to use in this research.

	Area of Use	Validity
MBTI	General model used in corporate settings	Weak
MMPI	Clinical model used in clinical and forensic settings	Strong
HEXACO	General model used in variety of studies	Strong
FFM	(Most common) General model widely used in variety of studies	Strong

Table 14 – Summary of area of use of personality models

Throughout the duration of the project, there was no evidence that the model of personality selected was the cause of any problems encountered. This selection led to a large design space that (while requiring a fair amount of interpretation) allowed for multiple design opportunities that supported the answering of SQ1 of the research question.

5.1.2 Questionnaire Selection

The second step in the methodology of this project was the selection of a questionnaire that was to be used as the basis of the designed video game tool. To this end, a survey of the literature was

undertaken in order to identify popular personality inventories within the model and subsequently selecting the one that was most suited for the project as described in Section 3.3.

In order to select a questionnaire to be referenced during the design process for the video game tool, three popular questionnaires were identified and analyzed (see Section 3.3). The BFI was rejected due to its lack of granularity as it was designed to be brief and report only at the factor level (see Section 3.3.1). The NEO-PI was rejected due to the cost associated with obtaining the questionnaire which not only covered monetary expenses, but also technical expertise (see Section 3.3.2). The IPIP was ultimately chosen because it provided instruction on its use and interpretation in an open source manner as well as offering questions which correspond to specific facets of the FFM which was intended to be useful in the design phase of the project (see Section 3.3.3).

Section 5.1.3.2 describes some of the difficulty experienced during the design phase with the options presented by the IPIP in terms of items that would serve as the basis for the video game tool's mechanics and metrics. Early in the design phase, the idea of using items from multiple questionnaire tools was entertained, but was rejected on the advice of subject matter experts (more detail in Section 5.1.3.1).

Firstly, as a minor consideration the IPIP itself shares a number of items with other contemporary questionnaires so that looking at a 300-item questionnaire wouldn't necessarily open up 300 items to consider anyway. The added variety would undoubtedly open avenues of design, but the primary reason for focusing singularly on the IPIP is that the different questionnaires all derive themselves from different mathematical distributions which could have a carry-on effect during the analysis phase.

On reflection, the IPIP served as a good starting point for the development of the video game tool. While the area of research that the project is situated in is still new, the wide spread usage of the questionnaire was useful for gaining an understanding of the typical contexts that it is used in. The open source nature of the tool also aided in the data analysis phase with clear and concise instruction on the scoring method for the questionnaire without the need for professional

interpretation. Based on the analysis that was carried out in this experiment, it may be worth considering the addition of other questionnaire's items into the design process in the future.

5.1.3 Deriving a Game Design from Questionnaire Items

Once the model of personality and associated questionnaire within that model were identified, the project moved to the task of generating game metrics and associated game mechanics from the aforementioned questionnaire (see Section 3.4). This task consisted of identifying promising behaviours to be measured (undertaken in consultation with subject matter experts), translating those behaviours to game metrics, and translating those metrics into mechanics to be implemented in the final game.

The following sub-sections will reflect on the process of selecting personality questions and deriving game mechanics.

5.1.3.1 Subject Matter Consultation and Initial Design

The first step in the process of deriving a game design from the personality questionnaire was the interpretation of items within the questionnaire (see Section 3.4.2). This ties back into answering SQ1 by first identifying the behaviours that the questionnaire measures with the goal of replicating it for the video game tool.

As this task relies on the interpretation of the questionnaire items, it is a relatively subjective task. In order to avoid misunderstandings, subject matter experts from the area of personality research were consulted: Professor Ken Kirkby, a professor of Psychiatry, and Dr. Allison Matthews from the division of Psychology within the School of Medicine and both from the University of Tasmania. The goal of this step was to identify the questionnaire items that could best be measured by in-game metrics. This would then form the basis of the answer to SQ1, charting out the opportunity space in designing a video game tool to measure personality and focusing specifically on the questionnaire and how that tool can be used in this context. This was done via a workshop that involved the subject matter experts from the fields of personality and video game design and relying on the experience of the experts to aid in the interpretation of the questionnaire items and translate them into behaviours to be observed (see Section 3.4.1).

The result was a list of behaviours which could be measured in a digital game environment which were then distilled into metrics to be measured within the video game as described in Section 3.4.2 and elaborated upon in Sections 3.4.2.1 to 3.4.2.4. The workshop process with subject matter experts included the requirements for each metric to be measured – if a metric worked only in a social system, only in the first encounter by the player, only in repeat encounters, etc. The workshop process also extended to what metrics would be feasible to collect for the project at hand, concerning time, budget, and alternative methods for ideal but impractical metrics. The goal was keeping the list as broad as possible at this stage to allow for the examination of as many facets as possible. This resulted in both the anonymous gift system (see Section 5.2.6) which promise for presenting social situations in a single player game as well as the aesthetic system (see Section 5.2.5) which showed the least player interaction.

The anonymous gift system and the results in Sections 0, 0, and 0 show that although the video game tool was a purely single player experience, it was able to measure social responses as outlined in Sections 3.4.2.3.1, 3.4.2.3.2, and 3.4.2.3.3. Conversely, this broad list also led to the implementation of the aesthetic system which was implemented to fulfill requirements outlined in Sections 3.4.2.2.2, 3.4.2.2.3, and 3.4.2.2.4 but which was ultimately ignored by a large portion of the players. The reasons for the low participation numbers in the aesthetic system as well as possible ways to improve the system to increase interaction with players is discussed in Section 5.2.5.

Both of the aforementioned systems could have been dismissed at this stage of the process; the anonymous gift system due to it being an inherently social system for a single player game, and the aesthetic system for how little it affected the game experience. Ultimately, the project benefitted more from the inclusion of those systems and would have lost useful data points if they had been dismissed here.

In hindsight, this process was necessarily vague in its execution given that the task of identifying game behaviours that can inform personality profile construction is still a new field with little prior work in the area (see Section 3.4). As a consequence, there were no guidelines to follow for

this undertaking, and any missteps taken during this project were necessary to the establishment of this process. In answering SQ1, this step has shown the value of subject matter expertise as well as the merit in keeping the possible design space broad at the onset of the project to facilitate more options later on.

5.1.3.2 Game Metrics to Game Mechanics

The next step was to derive the appropriate game mechanics that are required for the metrics identified and to combine them into a cohesive game design in order to answer SQ1. The goal here was to establish a game design that allowed for all of the needed metrics while still feeling like a sensible game design. This step continued to utilize the workshop format from before with the discussion focusing on which game mechanics would be needed in order to implement the game metrics previously identified while adhering to principles of good game design (Jesse Schell 2008; Despain et al. 2013; Salmond 2016) as described in Section 3.4.1.

The result of this step was the solidification of a game design as described in Sections 3.4.2.1 to 3.4.2.4. The process of arriving at that final game design was highly iterative and much of the mechanics that ultimately ended up in the video game were informed by the development and exploration of the video game itself. Extensive debates were held between subject matter experts during that process to determine how the game mechanics best fit together for a cohesive game while still allowing for the breadth of player behaviours to be measured in order to better understand the opportunities and challenges in the design process as described in SQ1.

However there still remained a number of FFM facets that were not operationalized into game mechanics. There are two main reasons why this remains the case even as the game was pushed to the experimentation phase. Firstly, certain facets were represented by highly subjective questionnaire items. Although design ideas for such items were theorized, most fell outside of the scope of the project either in time and/or budget. Secondly, the number of game mechanics currently implemented already pushed the game in many different directions. As the project stands at the time of writing of this document, the designed video game already tries to target as

much of the FFM facets as possible considering the experience level of the researchers involved in designing, implementing, and maintaining the video game tool.

Since the list of game metrics to be used to measure player behaviour was kept as broad as possible, this portion of the design process was forced to be spread wide across multiple in-game systems. Even metrics and mechanics that didn't make it into the final design were subject to extensive discussion as to the feasibility of their addition. Time constraints ultimately meant that the end of the possible allotted time for this portion of the design phase still saw promising metrics and associated mechanics under consideration.

One possible system that may have suffered from this compressed time frame is the NPC system as described in Sections 3.4.2.4.1 and 3.4.2.4.2 showing a low interaction rate in Section 0 which made data analysis unreliable. The specific reasons as to why this might be and how to potentially address this are discussed further in Section 3.5.7, but the NPC system in particular was added very late in the development cycle of the video game tool. As a result, it is likely that the system did not get the same level of iteration and scrutiny as other systems.

Thus, in answering SQ1, this process benefitted from subject matter experts from both the field of personality psychology and game design working with game design principles to determine the design of the video game tool. However, time limitations forced some systems such as the NPC system to undergo fewer design iterations which may have ultimately affected the rate at which players interacted with that system.

In the future, it is possible that a more rigid and regimented approach to the finalization of design would avoid neglected systems as described above. Specifying a hard cutoff period where no additional systems should be added to the game and focusing on iteration and testing of existing mechanics to that point would help ensure that all systems receive at least some care to their place in the flow of the game.

5.1.4 Lessons From Design

In summary, the design phase for this project in answering SQ1 granted many lessons to reflect on. Firstly, the area of personality research is filled with several models of personality that were

each developed to describe specific aspects of the human psyche. The FFM fulfilled this project's need for a general model and offered a large design space to answer the project's research question.

Secondly, even within the FFM of personality there exists a number of different tools for the purposes of eliciting personality information. The IPIP served as the best tool to be used as the basis for this project due to its open source nature which was useful in interpreting the questions within it.

Thirdly, the experience offered by subject matter experts in the fields of both personality research and game design proved to be invaluable in this phase. While a wealth of information regarding the interpretation of the personality questionnaire items is recorded in written form, the subject itself is still highly subjective and their guidance was important to avoid mistakes with regards to questionnaire item interpretation.

Fourthly, the decision to keep design options as broad as possible afforded more options as the project moved into development and potentially provided data that may have been overlooked. It is important to specify limitations to a project early on such as requiring it to be single player, but equally as important to keep an open mind when considering possible designs, such as how to adapt social mechanics into a single player experience.

Last but not least, the role that iteration should not be underestimated in projects such as these and therefore the schedule should reflect a suitable amount of time to allow for redesigns once development has begun in earnest.

5.2 Reflection on Implementation

This section will discuss the systems that were implemented in order to obtain the metrics needed to create a personality profile.

5.2.1 Inventory and Item System

The inventory and item system was implemented as described in the design specification that resulted from 3.4.2 and implemented in order to answer SQ2. For the facet of Self-Discipline, the

inventory and item system sought to give players the opportunity to make preparations for future obstacles as described in Section 3.4.2.1.5. For the facet of Orderliness, the inventory and item system sought to force the player to consider the organization of the items within their inventory as described in Section 3.4.2.1.1.

On reflection and in answering SQ2, the inventory and item system functioned mostly to the expectation of the design. Both systems forced players to utilize the systems in order to play the game at all. Levels within the game were constructed out of components that require item usage to traverse and the size of the inventory was kept purposefully small in order to force players to consider item management more consciously.

However, results from Sections 4.3.1 and 0 show that some players were able to fully traverse a level without using any items which means that players were effectively able to avoid the system altogether. This is an undesirable possibility for the system as it means that players can bypass the need to consider their item usage at all in some levels. The most likely reason for this is the role that gravity and falling plays in the traversal of the game world. While ascending and gaining height requires some degree of item manipulation, the random level generation may result in levels where the player starts at a higher elevation than the exit goal. If the level is generated in such a fashion, this would allow for the player to bypass many obstacles and fall towards the exit goal. In the future, this can be rectified with additional rules in the level generation system that forces the exit goal to be placed at a higher elevation than the starting point.

5.2.2 Scoring System

The scoring system was implemented as described in the design specification that resulted from Section 3.4.2 and implemented in order to answer SQ2. The scoring system was implemented as a number permanently displayed on-screen for the player as described in Section 3.4.2.1.4. This served to give players feedback on their progress through tasks that raise their score (such as the collection of coins). This also served to place a common language for players with experience in video games as scores are a common aspect of many other games (Jesse Schell 2008; Despain et al. 2013; Salmond 2016).

Ultimately no correlations were found with the metric to do with the scoring system as seen in Section 4.3.4. Some speculation as to why this might be the case (outside of other factors discussed later such as sample size) are that the score blends in too well in the game and is easily overlooked by players. Since there are currently only two actions that actively feed into the score – coin collection and anonymous gift interaction (which can only possibly occur once per level) – the player may quickly tune out the rising numbers. In the future, to make this system more noticeable to the player, sound effects can be used to support the feedback of the few actions that feed into the score.

On reflection, there is also the concern that the score implicitly tells the player that the game equates their collection of coins as a measure of their competency while internally the game is looking at the player's speed at completing levels (see Section 3.4.2.1.2). This contradiction has to be addressed and is a gap in the design that was overlooked by the steps discussed in Section 3.4. A possible solution for this is to integrate the time taken by the player as a component of the score system, granting additional points to this measure based on the speed at which players complete the level in addition to the coin collection aspect of the existing system.

A further method of highlighting the score as a measure of performance to the player and in theory creating more investment in the metric is discussed in Section 5.2.3.

5.2.3 Leaderboard System

The leaderboard system was implemented as described in the design specification that resulted from Section 3.4.2 and implemented in order to answer SQ2. The primary goals of the leaderboard was to allow players the opportunity to 'cheat' the system in order to assess how willing they were to break the game for the facet of Dutifulness (as described in Section 3.4.2.1.6) and to give players an opportunity to upload the results of their play session to a leaderboard for the facet of Modesty (as described in Section 3.4.2.3.4). The leaderboard was implemented as a randomly generated list of numbers that appears at the end of the level after the player has completed the puzzle. No identifying characters were associated with those numbers and the player's score (if they exceed any score in that list) is highlighted in yellow. The player was given the choice to

upload their score to the leaderboard should they desire. They were also given a text field that displayed the score that they had been given, allowing them to edit the uploaded score.

As shown in Sections 3.4.2.1.6 and 3.4.2.3.4, not every player managed to reach this point at all and thus a portion of the participants did not get to interact with this particular system. However, a much larger proportion of players (as shown in Section 0) managed to reach the exit point in the level but not complete the puzzle at the end. In the future, it is desirable to maximize the number of players interacting with the system in order to allow for better quality of data. To that end, it is possible to place the leaderboard upload mechanic at the point of the exit instead and offer players other incentives to complete the puzzle. This may include a bonus increase in score and a second opportunity to upload their score.

This system also showcases a potential to be used to reinforce the effectiveness of the Scoring system as described above (see Section 5.2.2). The leaderboard could be shown to the player at the beginning of a level in order to act as an incentive to pay attention to their score and to give the player some guidelines on the scores they would be required to beat in order to make it onto the leaderboard. Further improvements could see the leaderboard be populated with randomly generated names along with the scores (and subsequently allowing players to upload a name along with their score) in order to increase the feeling that the numbers shown on the leaderboard are genuine. However, such an addition should be done carefully so as to not repeat itself too often or being too transparently fake or it could run the risk of giving the impression of being a fake leaderboard.

5.2.4 Puzzle System

The puzzle system was implemented as described in the design specification that resulted from Section 3.4.2 and implemented in order to answer SQ2. A one screen block pushing puzzle was implemented as described in Sections 3.4.2.1.3 and 3.4.2.2.1. The puzzle revolves around pushing solidly coloured blocks onto correspondingly coloured goal positions in a grid (with full details on the specifics explained in Section 3.5.3). The puzzle was designed such that the entire puzzle space was viewable on one screen to allow the player to attempt solving the puzzle before even

making the first move. The game also recorded the player's actions and time spent as part of the assessment for the aforementioned metrics.

Sections 0 and 0 showed that most players did not start a puzzle and even among the ones who did, many did not finish it. It is possible that players felt uninterested in the puzzle, feeling the core experience of the game to be the exploration component. A possible reason for players not finishing a puzzle they start is how clearly the traversal and exploration aspect of the level is delineated from the puzzle component. Players may feel accomplished by reaching the exit at the 'end' of the level and decide not to go on once seeing what the puzzle has to offer.

A possible solution going forward for this to increase the amount of players who play the puzzle is to create a mid-point system that splits any given level into two parts that is separated by the puzzle, turning it into a mid-way checkpoint for the player. This would force players to complete the puzzle before moving onto the final section and attempting to reach the goal.

5.2.5 Aesthetic System

The aesthetic system was implemented as described in the design specification that resulted from Section 3.4.2 and implemented in order to answer SQ2. The goal was to give the player options that would allow them to select whichever aesthetic looked most appealing as described in Sections 3.4.2.2.2, 3.4.2.2.3, and 3.4.2.2.4. This system consisted of a menu outside of playing the game which would change the sprites that were loaded when the level was generated, thus changing the appearance of the level, the player's avatar, or both.

Section 0, 0, and 0 showed a low rate of use of the aesthetics options among participants of the experiment. The most likely explanation for this result is how separated the system is from the flow of the game, requiring the player to return to the main menu before being allowed to make aesthetic changes. Further, Section 0 showed a significant and strong negative correlation between the factor of Agreeableness and whether or not players chose to change the aesthetic of the game between levels. One possible reason for this is the low sample size causing a false positive result. As such, it is unclear if this is truly a significant correlation discovered by the game. However, in the possibility that it is, it can be explained by the notion that low

Agreeableness tends the player towards being dissatisfied with the look of the game. If this measure is expanded in the future, it is worth keeping this possible relationship in mind.

Section 0 showed a significant moderate negative correlation between the factor of Openness and whether or not players preferred the fantasy aesthetics over mundane aesthetics. Interestingly, this is an area in which a relationship was expected, however the theory and game design would have predicted a positive relationship. This result is also plagued by a small sample size and thus it is unclear if it truly is a significant correlation or a false positive result.

A possible method of increasing player interactivity with this system is to integrate the aesthetic changes more heavily into the mechanics of the game. As this would require drastic alterations to the game by way of introducing a new mechanic, careful thought should be given to this idea before it is implemented. One possible method for accomplishing this will be described here. Firstly, the player would need to be able to alter the aesthetic state of the game within a level, either by additional buttons in the UI, or a new menu option. Secondly, obstacles would need to be designed and placed in the level that required the use of changing aesthetics to solve. An example is a body of water that is impassable while liquid which can be frozen by changing to the snowy aesthetic (thus freezing the water to allow passage) or a wall of ice that impedes progress which must be transformed to the grass aesthetic to show it thawed and allow access.

Similar mechanics could be added onto the different character sprites by giving each specific version of the player avatar a distinct ability (perhaps jumping, crawling, or teleportation). It is important here that outside of specific obstacle events the player's avatar and the game world behave exactly the same regardless of aesthetic, thus allowing the player to choose freely between whichever options looked best to them (or even to make no choices outside of clearing obstacles).

5.2.6 Between-player Gift System

The gifting system was implemented as described in the design specification that resulted from Section 3.4.2 and implemented in order to answer SQ2. The goal of the system was to monitor player choice when it in an interaction with a pseudo-social system (as described in Sections 3.4.2.3.1 and 3.4.2.3.2) and functioned by placing a gift object close by to the player's spawn

position. When the player interacted with the object, they were given the option of accepting a gift from a random other player and told that it could be either a positive or negative gift. In reality, the decision on whether or not the effect is positive or negative was randomized, but this fact was hidden from the player in order to create a pseudo-social interaction.

A second aspect of the gift system was a reciprocal choice offered to the player upon reaching the exit. Players were asked if they would like to pass on a gift that was either positive or negative in nature to other random players. Players were able to ignore these options and simply start the puzzle without gifting anything. These player choices did not actually affect the gifts received by other players, but merely acted as a recording of their intent.

Sections 0 and 0 results show that there was a significant correlation between Agreeableness and either the gifting of positive a (weak positive correlation) or a negative gift (weak negative correlation). The results perfectly mirroring each other makes sense as most players who choose not to give a positive gift would give a negative one instead (and vice versa). While this is a hopeful result from the game, the sample size of occurrences of this decision (N = 28) is rather smaller than can be considered concrete as well as the sheer number of correlations that were tested open up the result to the possibility of a false positive. However, it is still worth noting this result in the future when the game can be exposed to a larger sample size for repeat and supporting results.

5.2.7 Non-playable Character (NPC) System

The non-playable character (NPC) system was implemented as described in the design specification that resulted from Section 3.4.2 and implemented in order to answer SQ2. The goal of this was to have the NPC groups serve as a pseudo-social group that the player may choose to gravitate towards or avoid at their own discretion as described in Sections 3.4.2.4.1 and 3.4.2.4.2. The system placed groups of NPCs around the level in locations that were not in the direct path to the exit. The player was also given a series of buttons that only activated at a certain distance away from an NPC group. These buttons displayed an 'emote' of varying emotions over the

player avatar's head. This emote was also replicated and displayed over the heads of any NPCs within a certain range. The game then recorded these button presses for analysis.

Section 0 showed that there was a significant moderate positive correlation between Neuroticism and the average number of 'Smile' emotes used by the player. The sample size for this correlation is small (N = 13) and the result could be a false positive with the amount of correlations that were tested. In the event that the result is accurate, a possible reason for the result could be that individuals who scored higher in Neuroticism sought deeper game mechanics within the simple 'emote' system presented and were experimenting with the mechanic. In the future, it would be advisable to pay close attention to this game mechanic and player's interaction with it.

In general, Sections 0 and 0 show that most players do not feel the need to interact with the NPC system. It is currently unclear how well these NPCs work as pseudo-social groups. Other pseudo-social mechanics in the game were hidden behind a layer of anonymity that allowed the game to assert human action behind random values, but these NPCs clearly do not serve such a purpose within the game world as they present the player with little of value. A possible avenue of making this mechanic more appealing to players and giving a role to NPCs is to create a system wherein 'other players' are able to implant NPCs with actions, messages, or items that are actually randomly generated by the game behind the scenes so as to create the sense of player interaction. A corresponding system would also have to be developed that allowed a player to implant those same actions, messages, or items to complete the illusion.

5.2.8 Participant Selection

The goal for participant selection was to obtain as general of a subset of the population as possible in order to test the game for the sake of answering SQ2. To this end, the recruitment criteria was described as opportunistic, aiming to obtain as large of a number of participants as possible to smooth along the data analysis. Invitations to the experiment were distributed on all available social media of the researchers involved as well as through mailing lists in the School of Engineering and ICT in the University of Tasmania.

This provided a good range of different demographics which tested the general appeal of the video game to a broad audience. However, the results in Section 4.3 show that this approach allows for inconsistencies between participants in terms of the amount of time spent with the game which in turn affects the interaction rates for every metric in the game. While this would be an acceptable and even expected result for a fully formed version of the video game tool, these results at this stage of the tool's development only serve to muddy the data analysis process, adding in more variables to the analysis. It is recommended that the participant recruitment be toned back down for the immediate future in order to pilot test the recommendations described in the previous sections and obtain a better understanding of the relationships between the metrics in the game and personality factors.

Conversely, once the video game tool itself is finalized, it is recommended to expand the recruitment pool in order to allow for more advanced regression analyses to be performed. While the methods used and described in Section 4.3 are adequate for determining correlates between game metrics and personality factors, more nuance can be extracted from the data given a larger sample size.

5.2.9 Procedure

The goal of the experimental procedure was to obtain a personality profile via traditional means (administration of a personality questionnaire) and compare those results to the results of the metrics measured by the game in fulfillment of SQ2. To that end, players were asked to play the game for as long as they wanted and to complete an IPIP questionnaire with the order of the two tasks being randomized to account for order effect.

While allowing the player to play for as long as they would like is an ideal way of administering the video game tool, this allowed for large variances in the play times as seen in Section 0 (which lead to potential large variances in game experience between players). For the purposes of the development of the video game as a tool, testing in the near future should be done under more lab-like conditions to allow for more uniform data gathering.

In the future, this procedure can also be improved by including a means of obtaining observer reports that would provide a third person report on the participant. As described in Section 2.2.2.4, observer reporting has the potential to address some of the deficiencies of social desirability bias. Due to the logistics of ensuring cooperation from multiple people per participant, it would be prudent to perform an initial test of the video game tool with observer report data in a more controlled setting first. A sensible approach would be to perform the experimentation under lab conditions with participants physically being present in order to ensure presence of data before developing the plan for larger scale general testing.

5.2.10 Lessons from Implementation

In reflection on the implementation process, there are many aspects of the designed video game tool that can stand to be tweaked or reworked in order to improve the system as a whole. However, there are also lessons that can be learned from this reflection on a more general note.

Firstly, systems that are designed in isolation may have unexpected effects on one another. For example, the method at which random level generation was executed affected results associated with the item and inventory system. While it is possible to identify these unexpected interactions at the design phase, it is often only revealed during testing.

Secondly, individual systems can be tied together to increase the cohesion of the game and help the game feel less like a bundle of systems and more like a complete game. With the way in which this project was carried out, many individual systems were designed and developed in isolation. Some thought and consideration can tie these systems together to create a more cohesive game with a simpler message for the player and increase engagement with the game.

Thirdly, the designed pace and flow of a game may not match the player experience. The designed video game tool had a very specific set of tasks for the player to accomplish, but results show that many players chose to end their session before reaching that point. It is important to step back and evaluate why something like this may have happened and then implement changes to help fix this problem in order to aid the data collection process.

Fourthly, in relation to the above it is likely that the beginning of a game will see more play than the end of a game as players spend more time trying to figure out the mechanics of a game. Additionally, due to the fact that games tend to be played voluntarily and for recreational purposes, players can feasibly be expected to simply stop playing at arbitrary points. If the in-situ nature of the project is important, then the design of the game should account for the important data to be collected as soon as possible to account for this phenomenon.

Fifthly, and in relation to the previous lesson, if it is important to have a complete play experience, projects such as these should consider lab conditions to enforce a specific amount of play. The control afforded by lab conditions especially in the earlier phases of development can help focus the project in on the data it is primarily concerned with before trying to manage extraneous conditions to the administration of the tool.

5.3 Summary

This chapter discussed the steps taken in the design phase as well as the system implementations of this project that were done in answering the research question:

> What are the opportunities and challenges in creating a serious game to support the identification of player personality characteristics?

In doing so, this chapter also highlighted several lessons learned along the process in relation to the two sub research questions which are presented in Table 15 below.

Sub Research Question	Lessons
1. What are the opportunities and challenges in *designing* a serious game to support the identification of player personality characteristics?	1. The FFM is a great general purpose model of personality that fits this project. 2. The IPIP's open source nature aided the design and development of this project. 3. Subject matter expertise can help with subjective interpretation.

	4. Broad initial designs help give the project options later on in development. 5. Ample time should be factored into the project for software iteration and redesign.
2. What are the opportunities and challenges in *implementing* a serious game to support the identification of player personality characteristics?	1. Separately designed systems may have unexpected effects on one another (random level generation affecting item usage). 2. Having systems reinforce messages of other systems helps to create a cohesive game rather than a collection of systems. 3. The perceived pace and flow of a game can affect where players stop playing. 4. The beginning of the game will see more play than the end of a game. Account for that in design if some data points are important. 5. While the value for in-situ experimentation cannot be underestimated, lab experiments give more control over variables which can aid data analysis.

Table 15 – Summary of lessons learned

The following chapter will present the conclusions achieved from this project, reflecting on the lessons learned and elaborating on the future direction of this project.

6. Conclusion

The research question of this project was defined as the following:

> What are the opportunities and challenges in creating a serious game to support the identification of player personality characteristics?

In order to answer it, two sub-research questions were developed that in turn answered the research question in parts. This chapter will focus on summarizing the answer to those questions and provide the final thoughts on the project.

6.1 Sub-Research Question 1

The first sub-research question is:

> What are the opportunities and challenges in *designing* a serious game to support the identification of player personality characteristics?

Looking specifically at the task of deriving a video game design from personality theory, Chapter 3 describes a multi-step progress. Firstly, Section 3.1 presents the literature and logic behind selecting the FFM as the model to base the project upon. This decision was carried out with a literature review and understanding of the extant models of personality. Section 3.3 followed with a selection of the best questionnaire – the IPIP – to be used for the project once more being based upon a literature review of the area, identifying the existing questionnaires and the specifics of their use. Using questionnaire items in the IPIP as the basis, 3.4 described the iterative process used to derive game design elements for the video game tool. The result of this entire process was a game design specification that was theoretically capable of observing player behaviours within a video game that would be used to predict personality information of the player.

Section 5.1 fully discusses the matter, but in summary and in answer of SQ-1, the design phase was a challenging undertaking owing primarily to the lack of established guidelines for the task of creating video game systems that would assess personality behaviour. This was further compounded by the fact that there was an initial lack of expertise within the area of personality

theory within the project team. In order to combat both problems, subject matter experts were consulted for their knowledge in the area. Even then there were challenges in balancing a large initial idea pool with limited time and resources needed to implement them. Ultimately the process yielded several lessons which are discussed in Section 5.1.4 which highlights the importance of keeping options open at the onset of the project, highly valuing the input of subject matter experts to aid interpretation of the theory, and emphasizing the importance of allowing for time to iterate over designs.

6.2 Sub-Research Question 2

The second sub-research question is:

> What are the opportunities and challenges in *implementing* a serious game to support the identification of player personality characteristics?

Section 3.5 describes the systems present in the video game tool in their final implemented forms. These systems were based off of the designs from the previous phase and each served the purpose of creating a space in which different the player's behaviour would be indicative of a particular facet of their personality. Section 3.6 then described the experiment that was carried out in order to test the efficacy of the video game tool which had a focus on in-situ responses.

In summary and in answer of SQ2, Sections 4.3 and 5.2 showed and reflected on the results of the video game tool in the experiment process. Sections 5.2.1 to 5.2.7 in particular discussed each system in detail and what the results could mean while Section 5.2.10 elaborated on the lessons learned from each system. The results were mixed, showcasing some of the major challenges of implementing a serious game system. Some of the systems (see Sections 5.2.3 and 5.2.4) showed low interaction rates which added a layer of complexity when interpreting the results. Other systems (see Sections 5.2.1 and 5.2.2) showed unexpected interactions between systems that may change the way players interact with the game which in turn adds unexpected variables to the interpretation of results. Not all systems presented problems though, with some systems (see Sections 5.2.3, 5.2.6, and 5.2.7) showing potential at creating pseudo-social scenarios in a single player game experience. This is useful to allow the video game tool to create more varied

scenarios for players without the overhead of more complex social interaction systems within the game itself. As Sections 5.2.1 to 5.2.9 discussed the results of each system and speculated on why any unexpected results may have occurred, potential improvements to each system was also elaborated with the goal of improving the video game and making it seem more like a cohesive game as opposed to a loose collection of systems.

6.3 Future Work

The results obtained from Chapter 4 show that the game is not yet able to accurately predict personality measures. Section 5.2 discussed modifications that can be made to the video game tool's systems to improve the video game, specifically targeting the rate at which players interacted with systems and the cohesiveness of the game's many systems to deliver a stronger player experience.

Most of the suggested changes are small in scope, such as adding extra heuristics to the level generation to enforce item usage (see Section 5.2.1) and discussed more completely in their relevant sections between Sections 5.2.1 and 5.2.8. Some other suggestions cover a larger change to the system such as the expansion of the aesthetics system (see Section 5.2.5) to integrate new functionality into the game that would give the different aesthetics changes an added purpose. Some of the game's systems could also stand to be iterated upon to increase cohesion within the game and present the player with a more uniform message. A notable example of this would be the scoring and leaderboard systems (see Sections 5.2.2 and 5.2.3) which both serve the purposes of providing feedback to the player on their performance but are currently treated as separate systems that barely interact with one another.

Another avenue that has to be taken into consideration for the future of this project or projects like this are the implications of the context in which player behaviour is being recorded and analyzed as discussed in Section 2.3. Although it was too late to incorporate lessons from Canossa et al.'s (2015) research, it would undoubtedly have an effect on the design of the video game tool, potentially pushing towards a more separated and segmented game experience with smaller foci

on specific player behaviours as opposed to a large overarching game system that targets many facets at once.

The experiment that was already conducted aimed to obtain in-situ data from a general subset of the population, but the results obtained (see Section 4.3) showed a much larger variance than is helpful for testing at this stage. Thus, future testing sessions were discussed in Sections 5.2.8 and 5.2.9 to consider more lab-like conditions to first focus on streamlining the game flow and overall experience. Further suggestions also include the expansion of the experimental procedure to include tertiary observer reports to serve as an additional personality data point for analysis.

Overall this experiment showed that the designed video game tool has promise, but is still in an early state of design before being ready for the large scale testing that would be required to truly compare the results of this tool with traditional methods.

6.4 Primary Research Question

In answering the two sub-research questions, the primary research question is now able to be answered. The major obstacle and challenge to the creation of a serious game that supports the identification of player personality characteristics is the lack of precedence in the area. While literature shows work done with existing video games and some attempts had been made at custom designing scenarios to test for player personality (see Section 2.2.2.6.3), there were no guidelines or precedence when it came to a behaviourally focused video game tool. This was primarily overcome with knowledge provided by subject matter experts and followed the steps outlined in Section 3.4 to create the video game tool's design. Following this, an experiment was conducted with the results showing that not all of the designs performed according to predictions, but also demonstrating some interesting relationships. While the tool itself is not yet ready to be used as a new method of personality data elicitation, this project also describes improvements to be made to the existing software to help in that regard.

6.5 Contributions

This section describes the contributions that this work makes to theory, practice, and methodology.

6.5.1 Contribution to Theory

Prior to this work, the only examples of serious games that aimed to assess personality functioned primarily as a vehicle for delivering multiple choice questions to the player (Van Lankveld, Spronck, et al. 2011; Van Lankveld et al. 2009). By creating a serious game that observes player behaviour as opposed to getting questions answered, this work has contributed a way of approaching the design process of a serious game that is informed by the underlying theory it is assessing.

This work has contributed knowledge on the nuances and challenges of designing a serious game that is based on observing player behaviour, and in doing so contributed to the body of knowledge in the fields of personality psychology and serious games.

This thesis contributes to the ever-growing body of work regarding the assessment of personality which may potentially have the ability to predict health problems such as depression and anxiety (Lewis et al. 2014; Chow & Roberts 2014), cardiovascular risk (Gleason et al. 2014), metabolism (Human et al. 2013; Israel et al. 2014), and mental disorders (Trull 2012; Mullins-Sweatt & Lengel 2012; Widiger & Presnall 2013).

6.5.2 Contribution to Practice

This work has contributed to practice by presenting an example of a serious game created with the intent of measuring personality behaviour that is primarily behavioural. Further, the challenges and lessons learned from this process contributes to the relatively new field of designing serious games that are embedded in the theories they support, giving future researchers an idea of the things they should avoid and consider.

The designed video game tool is a serious game with an emphasis on offering traditional video game entertainment and performing its behavioural observation surreptitiously. The game contributes to a growing body of work in the area of serious games with strong production values such as the studies by Bindoff et al. (2014).

6.5.3 Contribution to Methodology

This work has contributed to methodology by presenting a new approach to creating a serious game that incorporates personality theory into the design process. The steps taken in this project (described in Chapter 3) favour a focus on the knowledge provided by subject matter experts to guide the design phase of the project. It attempts to address the gap in knowledge regarding the best methods of deriving game design from an established theory with an in-depth discussion process as well as iterating over promising designs. These processes were also reviewed in hindsight with the results of the experiment to recommend improvements to the process described in Section 5.1.

The systems of the video game tool are also discussed in reflection with the results of the experiment in mind in order to generate a list of recommendations that would improve the cohesiveness and performance of the video game tool (see Section 5.2).

In the cases of the discussion of the design and implementation phase of the video game tool, the lessons learned were also distilled into more general statements that may benefit future researchers with the goal of creating serious games (see Section 5.3).

7. References

Abe, J.A. a., 2005. The predictive validity of the Five-Factor Model of personality with preschool age children: A nine year follow-up study. *Journal of Research in Personality*, 39(4), pp.423–442. Available at: http://linkinghub.elsevier.com/retrieve/pii/S009265660400039X [Accessed November 7, 2014].

Abt, C.C., 1970. *Serious Games*, Viking Press. Available at: https://books.google.com.au/books/about/Serious_Games.html?id=axUs9HA-hF8C&printsec=frontcover&source=kp_read_button&redir_esc=y#v=onepage&q&f=false.

von Ahn, L. & Dabbish, L., 2008. Designing games with a purpose. *Communications of the ACM*, 51(8), p.57.

von Ahn, L. & Dabbish, L., 2004a. ESP Game. Available at: https://crowdsourcingpower.wordpress.com/2010/01/28/esp-game-or-google-image-labeler-great-ideas-based-on-crowdsourcing/ [Accessed November 3, 2017].

von Ahn, L. & Dabbish, L., 2004b. Labeling images with a computer game. *Proceedings of the 2004 conference on Human factors in computing systems - CHI '04*, pp.319–326. Available at: http://portal.acm.org/citation.cfm?doid=985692.985733.

Allport, G.W. & Odbert, H.S., 1936. Trait Names: A Psycho-lexical Study. *Psychological Monographs*, 47, p.171. Available at: http://psych.colorado.edu/~carey/courses/psyc5112/Readings/psnTraitNames_Allport.pdf.

APA, 2015. American Psychological Association. Available at: http://www.apa.org/topics/personality/.

Archer, R.P. et al., 2006. A survey of psychological test use patterns among forensic psychologists. *Journal of personality assessment*, 87(1), pp.84–94. Available at: http://ikpp.si/att/29/Archer - Survey of forensic test usage.pdf.

Arney, F., 2004. A comparison of direct observation and self-report measures of parenting

behaviour. Available at: http://ebooks.adelaide.edu.au/dspace/handle/2440/37713.

Bai, S. et al., 2013. Predicting Big Five Personality Traits of Microblog Users. In *Web Intelligence (WI) and Intelligent Agent Technologies (IAT)*. Available at: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6690057 [Accessed September 17, 2014].

Ben-Porath, Y.S., 2013. Using the MMPI-2-RF validity scales in forensic assessments. *The Gavel*. Available at: http://www.apadivisions.org/division-18/publications/newsletters/gavel/2013/07/forensic-assessments.aspx [Accessed January 31, 2015].

Bethesda, 2009. Fallout 3. *Bethesda Softworks LLC*. Available at: http://store.steampowered.com/app/22370/Fallout_3_Game_of_the_Year_Edition/.

Bindoff, I. et al., 2014. INSTRUCTIONAL DESIGN AND ASSESSMENT A Computer Simulation of Community Pharmacy Practice for Educational Use. *American Journal of Pharmaceutical Education*, 78(9), pp.1–8.

Bindoff, I. et al., 2016. Quittr: The Design of a Video Game to Support Smoking Cessation. *JMIR Serious Games*, 4(2), p.e19. Available at: http://games.jmir.org/2016/2/e19/.

BioWare, 2002. Neverwinter Nights. Available at: http://au.ign.com/wikis/neverwinter-nights.

Boldero, J., 2007. Convergence between GNAT assessed implicit and explicit personality. *European Journal of Personality*, 358(December 2006), pp.341–358. Available at: http://onlinelibrary.wiley.com/doi/10.1002/per.622/abstract [Accessed September 17, 2014].

Bond, M.H., Nakazato, H. & Shiraishi, D., 1975. Universality and Distinctiveness in Dimensions of Japanese Person Perception. *Journal of Cross-Cultural Psychology*, 6(3), pp.346–357. Available at: http://jcc.sagepub.com/content/6/3/346.short.

Borkenau, P. & Ostendorf, F., 1990. Comparing exploratory and confirmatory factor analysis: A study on the 5-factor model of personality. *Personality and Individual Differences*, 11, pp.515–

524. Available at: http://citeseerx.ist.psu.edu/viewdoc/download;jsessionid=B884593D4D8A8D3C369F84A0E61CAA9B?doi=10.1.1.453.4286&rep=rep1&type=pdf.

Boyle, G.J., 1995. Myers-Briggs Type Indicator (MBTI): Some Psychometric Limitations. *Review Literature And Arts Of The Americas*, 30, pp.71–71. Available at: http://epublications.bond.edu.au/cgi/viewcontent.cgi?article=1026&context=hss_pubs.

Brewer, R.S., Lee, G.E. & Johnson, P.M., 2011. The Kukui Cup: A dorm energy competition focused on sustainable behavior change and energy literacy. *Proceedings of the Annual Hawaii International Conference on System Sciences*.

Burnett, D., 2013. Nothing personal: The questionable Myers-Briggs test. *Guardian News and Media Limited2*. Available at: http://www.theguardian.com/science/brain-flapping/2013/mar/19/myers-briggs-test-unscientific.

Butcher, J.N. et al., 1190. *Development and use of the MMPI-2 Content Scales. MMPI-2 monograph series.*, Minneapolis: University of Minnesota Press. Available at: http://psycnet.apa.org/psycinfo/1990-97427-000.

Butcher, J.N., Ones, D.S. & Cullen, M., 2006. Personnel Screening With the MMPI-2. In J. N. Butcher, ed. *MMPI-2: A practitioner's guide*. pp. 381–406. Available at: http://psycnet.apa.org/books/11287/014.

Butcher, J.N. & Williams, C.L., 2009. Personality Assessment with the MMPI-2: Historical Roots, International Adaptations, and Current Challenges. *Applied Psychology Health and WellBeing*, 1(1), pp.105–135. Available at: http://blackwell-synergy.com/doi/abs/10.1111/j.1758-0854.2008.01007.x.

Camara, W.J., Nathan, J.S. & Puente, A.E., 2000. Psychological test usage: Implications in professional psychology. *Professional Psychology: Research and Practice*, 31(2), pp.141–154. Available at: http://antonioepuente.com/wp-content/uploads/2013/01/2000.Camara-

Nathan-and-Puente-2000-psych-test-usage.pdf.

Canossa, A. et al., 2015. In Your Face(t) Impact of Personality and Context on Gameplay Behavior. *Foundations of Digital Games*, (September 2016). Available at: https://pdfs.semanticscholar.org/9041/3435cbd55e0f86f1e103f4ac55eefab8bf39.pdf.

Canossa, A., Martinez, J.B. & Togelius, J., 2013. Give me a reason to dig Minecraft and psychology of motivation. In *Computational Intelligence in Games (CIG)*. Niagara Falls, ON: IEEE, pp. 1–8. Available at: http://julian.togelius.com/Canossa2013Give.pdf.

Cantador, I., Fernández-tobías, I. & Bellogín, A., 2013. Relating Personality Types with User Preferences in Multiple Entertainment Domains. *EMPIRE 1st Workshop on "Emotions and Personality in Personalized Services", 10. June 2013, Rome*. Available at: http://ceur-ws.org/Vol-997/empire2013_paper_2.pdf.

Cape, P., 2010. Questionnaire length, fatigue effects and response quality revisited. *Survey Sampling International*, (August). Available at: http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Questionnaire+Length+,+Fatigue+Effects+and+Response+Quality+Revisited#0 [Accessed September 21, 2014].

Chamorro-Premuzic, T., Furnham, A. & Moutafi, J., 2004. The relationship between estimated and psychometric personality and intelligence scores. *Journal of Research in Personality*, 38(5), pp.505–513. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0092656603001119 [Accessed November 6, 2014].

Chen, S.X., Benet-Martínez, V. & Ng, J.C.K., 2014. Does language affect personality perception? A functional approach to testing the Whorfian hypothesis. *Journal of personality*, 82(2), pp.130–43. Available at: http://www.ncbi.nlm.nih.gov/pubmed/23607801 [Accessed November 13, 2014].

Cherry, K., 2015a. The 10 Scales of the MMPI. *About.com*. Available at: http://psychology.about.com/od/psychologicaltesting/a/mmpi_2.htm [Accessed January 30, 2015].

Cherry, K., 2015b. The Big Five Personality Dimensions. *About.com*. Available at: http://psychology.about.com/od/personalitydevelopment/a/bigfive.htm.

Chittaranjan, G., Blom, J. & Gatica-Perez, D., 2013. Mining large-scale smartphone data for personality studies. *Personal and Ubiquitous Computing*, 17(3), pp.433–450. Available at: http://link.springer.com/10.1007/s00779-011-0490-1 [Accessed September 17, 2014].

Chow, P.I. & Roberts, B.W., 2014. Examining the relationship between changes in personality and changes in depression. *Journal of Research in Personality*, 51, pp.38–46. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0092656614000439 [Accessed December 4, 2014].

Connelly, B.S. & Hülsheger, U.R., 2012. A narrower scope or a clearer lens for personality? Examining sources of observers' advantages over self-reports for predicting performance. *Journal of personality*, 80(3), pp.603–31. Available at: http://www.ncbi.nlm.nih.gov/pubmed/22091626 [Accessed December 9, 2014].

Cooper, S. et al., 2011. Analysis of Social Gameplay Macros in the Foldit Cookbook. In *6th International Conference on Foundations of Digital Games*. pp. 9–14. Available at: https://dl.acm.org/citation.cfm?id=2159367.

Cooper, S. et al., 2008. foldit. Available at: https://fold.it/portal/ [Accessed November 3, 2017].

Costa, P.T. & McCrae, R.R., 2016. ACER Shop. Available at: https://shop.acer.edu.au/neo-personality-inventory-revised-neo-pi-r.

Costa, P.T. & McCrae, R.R., 1992. *Revised NEO Personality Inventory (NEO PI-R) and NEO Five-Factor Inventory (NEO-FFI)*,

Costa, P.T. & McCrae, R.R., 1985. *The NEO Personality Inventory: Manual, Form S and Form R*,

Costa, P.T.J. & McCrae, R.R., 1976. Age Differences in Personality Structure: a Cluster Analytic Approach. *Journal of Gerontology*, 31(5), pp.564–570. Available at: http://geronj.oxfordjournals.org.ezproxy.utas.edu.au/content/31/5/564.abstract.

Costa, P.T.J. & McCrae, R.R., 2008. *The Revised NEO Personality Inventory (NEO-PI-R). The SAGE Handbook of Personality Theory and Assessment: Volume 2 — Personality Measurement and Testing. SAGE Publications Ltd*, SAGE Publications Ltd. Available at: http://dx.doi.org/10.4135/9781849200479.

Crowne, D.P. & Marlowe, D., 1960. A NEW SCALE OF SOCIAL DESIRABILITY INDEPENDENT OF PSYCHOPATHOLOGY. *Journal of consulting psychology*, 24(4), pp.349–354.

Dattner, B., 2008. The use and misuse of personality tests for coaching and development. *Sussex Publishers*. Available at: https://www.psychologytoday.com/blog/credit-and-blame-work/200806/the-use-and-misuse-personality-tests-coaching-and-development [Accessed January 28, 2015].

DeGraft-Johnson, C. et al., 2013. *Relating Five Factor Personality Traits to Video Game Preference*, Available at: http://hcil2.cs.umd.edu/trs/2013-08/2013-08.pdf.

Despain, W. et al., 2013. *100 Principles of Game Design*,

Deterding, S. et al., 2011. From game design elements to gamefulness. *Proceedings of the 15th International Academic MindTrek Conference on Envisioning Future Media Environments - MindTrek '11*, pp.9–11. Available at: http://doi.acm.org/10.1145/2181037.2181040%255Cnhttp://dl.acm.org/citation.cfm?doid=2181037.2181040.

DeYoung, C.G. et al., 2014. Openness to experience, intellect, and cognitive ability. *Journal of personality assessment*, 96(1), pp.46–52. Available at: http://www.ncbi.nlm.nih.gov/pubmed/23795918 [Accessed November 20, 2014].

Digman, J.M. & Takemoto-Chock, N.K., 1981. Factors In The Natural Language Of Personality: Re-Analysis, Comparison, And Interpretation Of Six Major Studies. *Multivariate Behavioral Research*, 16(2), pp.149–170. Available at: http://www.tandfonline.com.ezproxy.utas.edu.au/doi/abs/10.1207/s15327906mbr1602_2.

Dobewall, H. et al., 2014. A comparison of self-other agreement in personal values versus the Big Five personality traits. *Journal of Research in Personality*, 50, pp.1–10. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0092656614000051 [Accessed December 4, 2014].

Egenfeldt-Nielsen, S., Smith, J.H. & Tosca, S.P., 2009. *Understanding Video Games: The Essential Introduction*, Routledge. Available at: https://books.google.com.au/books?id=JYKQAgAAQBAJ&dq=video+game+industry+&lr.

Entertainment Software Association, 2017. Essential facts 2017. *ESA Report 2017*, p.20. Available at: http://www.theesa.com/wp-content/uploads/2017/04/EF2017_FinalDigital.pdf.

Eysenck, H.J. & Eysenck, S.B.G., 1975. *Manual of the Eysenck Personality Questionnaire (junior and Adult)*, Hodder and Stoughton. Available at: http://books.google.com.au/books/about/Manual_of_the_Eysenck_Personality_Questi.html?id=_uXrNAAACAAJ&redir_esc=y.

Ezzo, F.R., Pinsoneault, T.B. & Evans, T.M., 2008. A Comparison of MMPI-2 Profiles Between Child Maltreatment Cases and Two Types of Custody Cases. *Journal of Forensic Psychology Practice*, (February 2015), pp.37–41. Available at: http://dx.doi.org/10.1300/J158v07n02_02 PLEASE.

Fisher, R.J., 1993. Social Indirect Desirability Questioning Bias and the Validity of Indirect Questioning. *Journal of Consumer Research, Inc.*, 20(2), pp.303–315.

Furler, K., Gomez, V. & Grob, A., 2014. Personality perceptions and relationship satisfaction in couples. *Journal of Research in Personality*, 50, pp.33–41. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0092656614000178 [Accessed December 1, 2014].

FVDisco, 2011. Minecraft Programmable Piano. Available at: https://www.youtube.com/watch?v=qt1fOED0vC8 [Accessed March 5, 2015].

Geelan, B. et al., 2016. Augmented exergaming. *Proceedings of the 28th Australian Conference on Computer-Human Interaction - OzCHI '16*, pp.542–551. Available at: http://dl.acm.org/citation.cfm?doid=3010915.3010940.

Gleason, M.E.J. et al., 2014. The enduring impact of maladaptive personality traits on relationship quality and health in later life. *Journal of personality*, 82(6), pp.493–501. Available at: http://www.ncbi.nlm.nih.gov/pubmed/23998798 [Accessed November 13, 2014].

Gnambs, T., 2014. A meta-analysis of dependability coefficients (test–retest reliabilities) for measures of the Big Five. *Journal of Research in Personality*, 52, pp.20–28. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0092656614000543 [Accessed November 21, 2014].

Goldberg, L.R., 1971. A historical survey of personality scales and inventory. In P. McReynolds, ed. *Advances in Psychological Assessment*. Science and Behavior Books, Inc., pp. 293–336. Available at: http://projects.ori.org/lrg/PDFs_papers/Goldberg_1971_Historical_Survey.pdf.

Goldberg, L.R., 1999. *International Personality Item Pool: A Scientific Collaboratory for the Development of Advanced Measures of Personality and Other Individual Differences*,

Goldberg, L.R., 1981. Language and Individual Differences: The Search for Universals in Personality Lexicons. *Review of Personality and Social Psychology*, 2, pp.141–165. Available at: http://projects.ori.org/lrg/PDFs_papers/universals.lexicon.81.pdf.

Gordon, R.M., 2011. Definitions of MMPI/MMPI-2. Available at: http://www.mmpi-info.com/mmpi-2 [Accessed January 30, 2015].

Greene, R.L., 1990. *The MMPI-2: An interpretive manual (2nd ed.).*,

Grimm, P., 2010. *Social Desirability Bias*,

Gurven, M. & Rueden, C. von, 2013. How universal is the Big Five? Testing the five-factor model of personality variation among forager–farmers in the Bolivian Amazon. *Journal of personality*

and social psychology, 104(2), pp.354–70. Available at: http://www.ncbi.nlm.nih.gov/pubmed/23245291 [Accessed September 17, 2014].

Harkness, A.R. et al., 2014. The MMPI-2-RF Personality Psychopathology Five (PSY-5-RF) scales: development and validity research. *Journal of personality assessment*, 96(2), pp.140–50. Available at: http://www.ncbi.nlm.nih.gov/pubmed/23941166 [Accessed November 22, 2014].

Hilbig, B.E., Glöckner, A. & Zettler, I., 2014. Personality and prosocial behavior: linking basic traits and social value orientations. *Journal of personality and social psychology*, 107(3), pp.529–39. Available at: http://www.ncbi.nlm.nih.gov/pubmed/25019254 [Accessed November 10, 2014].

Human, L.J. et al., 2013. Is change bad? Personality change is associated with poorer psychological health and greater metabolic syndrome in midlife. *Journal of personality*, 81(3), pp.249–60. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3655908&tool=pmcentrez&rendertype=abstract [Accessed December 7, 2014].

Human, L.J. et al., 2014. To thine own self be true: psychological adjustment promotes judgeability via personality-behavior congruence. *Journal of personality and social psychology*, 106(2), pp.286–303. Available at: http://www.ncbi.nlm.nih.gov/pubmed/24467423 [Accessed November 9, 2014].

Hunter, H.K. et al., 2014. Using the MMPI-2-RF to Discriminate Psychometrically Identified Schizotypic College Students From a Matched Comparison Sample. *Journal of personality assessment*, 96(6), pp.596–603. Available at: http://www.ncbi.nlm.nih.gov/pubmed/24906115 [Accessed November 22, 2014].

IGEA, 2017. Taking it up a level: Australian video game industry generates $2.958B in 2016. *IGEA*. Available at: http://www.igea.net/2017/02/taking-in-up-a-level-australian-video-game-industry-generates-2-958b-in-2016/ [Accessed November 1, 2017].

Inc., B.E., 2015. World of Warcraft. *Blizzard Entertainment Inc.* Available at: http://us.battle.net/wow/en/ [Accessed February 6, 2015].

Inc., S.P.S., 2002. America's Army. *San Pedro Software Inc.* Available at: http://www.old-games.com/download/6796/america-s-army [Accessed November 3, 2017].

Isachsen, O. & Berens, L. V., 1998. *Working Together* 3rd ed., San Juan Capistrano: Institute for Management Development.

Israel, S. et al., 2014. Translating personality psychology to help personalize preventive medicine for young adult patients. *Journal of personality and social psychology*, 106(3), pp.484–98. Available at: http://www.ncbi.nlm.nih.gov/pubmed/24588093 [Accessed November 16, 2014].

Jaakko, S., 2012. In Defence of a Magic Circle: The Social and Mental Boundaries of Play. In *DiGRA Nordic '12: Proceedings of 2012 International DiGRA Nordic Conference*. Available at: http://www.digra.org/digital-library/publications/in-defence-of-a-magic-circle-the-social-and-mental-boundaries-of-play/ [Accessed January 19, 2016].

Jesse Schell, 2008. *The art of game design*,

John, O.P., Donahue, E.M. & Kentle, R.L., 1991. *The Big Five Inventory - Versions 4a and 54*,

John, O.P., Goldberg, L.R. & Angleitner, A., 1984. Better than the alphabet: Taxonomies of personality-descriptive terms in English, Dutch, and German. *Personality psychology in Europe, Vol.1: Theoretical and empirical developments*, pp.83–100. Available at: http://pub.uni-bielefeld.de/download/1779419/2312704.

John, O.P. & Srivastava, S., 1999. BIG FIVE INVENTORY (BFI). , 2.

John, O.P. & Srivastava, S., 1999. The Big Five trait taxonomy: History, measurement, and theoretical perspectives. *Handbook of personality: Theory and research*, 2(510), pp.102–138. Available at: http://books.google.com/books?hl=en&lr=&id=b0yalwi1HDMC&oi=fnd&pg=PA102&dq=T

he+big-five+trait+taxonomy:+History,+Measurement,+and+Theoretical+Perspectives.&ots=756zS6ZtPk&sig=-3pfI7eNKlyZLlJYEmwdDYeJ82Y%5Cnhttp://scholar.google.de/scholar?hl=de&q=john+sriva.

Johnson, J. a., 2014. Measuring thirty facets of the Five Factor Model with a 120-item public domain inventory: Development of the IPIP-NEO-120. *Journal of Research in Personality*, 51, pp.78–89. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0092656614000506 [Accessed December 4, 2014].

Johnson, J.A., 2008. *The IPIP-NEO Personality Assessment Tools*,

Kazdin, A.E., 2000. *Encyclopedia of Psychology*,

Khatib, F. et al., 2011. Crystal Structure of a Monomeric Retroviral Protease Solved by Protein Folding Game Players. *Nature Structural and Molecular Biology*, 18(10), pp.1175–1177. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3705907/.

King, M.F. & Bruner, G.C., 2000. Social desirability bias: A neglected aspect of validity testing. *Psychology & Marketing*, 17(2), pp.79–103.

Kirkby, K.C. et al., 1999. Behavioral Analysis of Computer-Administered Vicarious Exposure in Agoraphobic Subjects: The Effect of Personality on In-Session Treatment Process. *Comprehensive Psychiatry*, 40(5), pp.386–390. Available at: http://www.ncbi.nlm.nih.gov/pubmed/10509622.

Klimstra, T. a. et al., 2014. The Dark Triad of personality in adolescence: Psychometric properties of a concise measure and associations with adolescent adjustment from a multi-informant perspective. *Journal of Research in Personality*, 53, pp.84–92. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0092656614000981 [Accessed November 21, 2014].

Van Lankveld, G., Schreurs, S., et al., 2011. Extraversion in games. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 6515 LNCS, pp.263–275.

Van Lankveld, G., Spronck, P., et al., 2011. Games as personality profiling tools. *2011 IEEE Conference on Computational Intelligence and Games, CIG 2011*, pp.197–202.

Van Lankveld, G., Schreurs, S. & Spronck, P., 2009. Psychologically verified player modelling. *Gameon*, pp.12–19. Available at: http://www.gielvanlankveld.nl/literature/Psychologically Verifed Player Modelling.pdf.

Lee, K. & Ashton, M.C., 2015. The HEXACO Personality Inventory - Revised. Available at: http://hexaco.org/scaledescriptions [Accessed February 2, 2015].

Letzring, T.D. & Human, L.J., 2014. An examination of information quality as a moderator of accurate personality judgment. *Journal of Personality*, 82(5), pp.440–51. Available at: http://www.ncbi.nlm.nih.gov/pubmed/24118039 [Accessed November 13, 2014].

Lewis, G.J. et al., 2014. Core dimensions of personality broadly account for the link from perceived social support to symptoms of depression and anxiety. *Journal of Personality*, 82(4), pp.329–39. Available at: http://www.ncbi.nlm.nih.gov/pubmed/23957560 [Accessed November 13, 2014].

Liu, D., Li, X. & Santhanam, R., 2013. Research Article Digital G Ames and Beyond : What Happens When Players Compete? *MIS Quarterly*, 37(1), pp.111–124.

Livingston, R.B. et al., 2006. MMPI--2 Code-Type Congruence of Injured Workers. *Psychological Assessment*, 18(1), pp.126–130. Available at: http://psycnet.apa.org/journals/pas/18/1/126/.

Lofgren, E.T. & Fefferman, N.H., 2007. The untapped potential of virtual game worlds to shed light on real world epidemics. *The Lancet. Infectious diseases*, 7(9), pp.625–9. Available at: http://www.sciencedirect.com/science/article/pii/S1473309907702128 [Accessed March 16, 2015].

Loo, R. & Thorpe, K., 2000. Confirmatory factor analyses of the full and short versions of the Marlowe-Crowne Social Desirability Scale. *The Journal of Social Psychology*. Available at: http://www.tandfonline.com/doi/abs/10.1080/00224540009600503 [Accessed November 15, 2014].

Marsh, T., 2011. Serious games continuum : Between games for purpose and experiential environments for purpose. , 2, pp.61–68.

Mason, M., 2014. Pirate Metrics: How to track + improve your games core metrics. *Magmic Inc.* Available at: http://developers.magmic.com/metrics-track-mobile-game/ [Accessed March 10, 2015].

McCrae, R.R. et al., 2010. The Validity and Structure of Culture-Level Personality Scores: Data From Ratings of Young Adolescents. *Journal of Personality*, 78(June), pp.815–838. Available at: http://onlinelibrary.wiley.com/doi/10.1111/j.1467-6494.2010.00634.x/abstract.

McCrae, R.R. & Costa, P.T.J., 2012. *Personality in Adulthood: A Five-factor Theory Perspective* 2nd ed., Guilford Press. Available at: http://books.google.com.au/books?id=cBxSiWT_fSgC&source=gbs_book_other_versions.

McCrae, R.R. & Costa, P.T.J., 1989. Reinterpreting the Myers-Briggs Type Indicator from the perspective of the five-factor model of personality. *Journal of personality*, 57(March 1989), pp.17–40.

McCrae, R.R. & John, O.P., 1992. An introduction to the five factor model and its applications. *Journal of Personality*, 60(2), pp.175–215. Available at: http://www.ncbi.nlm.nih.gov/pubmed/1635039 [Accessed September 17, 2014].

McCrae, R.R. & Terracciano, A., 2005. Universal Features of Personality Traits From the Observer's Perspective: Data From 50 Cultures. *Journal of Personality and Social Psychology*, pp.547–561.

McLeroy, C., 2008. Improving "America's Army." *Soldiers*, 63(9), pp.7–9. Available at:

http://www.dtic.mil/dtic/aulimp/citations/gsa/2008_149928/169596.html.

Michael, D.R. & Chen, S.L., 2006. *Serious Games: Games that Educate, Train, and Inform*,

Microsoft, 2015. Xbox 360. *Microsoft*. Available at: http://www.xbox.com/en-AU/Xbox360/index.

Moore, M.E., 2011. *Basics of game design*, A K Peters/CRC Press. Available at: https://books.google.com.au/books?id=E9JG6JjPU-sC&dq=game+design&lr=&source=gbs_navlinks_s [Accessed April 20, 2017].

Mortel, T. Van de, 2008. Faking it: social desirability response bias in self-report research. *Australian Journal of Advanced Nursing*, 25(4), pp.40–48. Available at: http://search.informit.com.au/documentSummary;dn=210155003844269;res=IELHEA [Accessed October 12, 2014].

Moss, S., 2008. Crowne and Marlowe Social Desirability Scale. *Psychlopedia*.

Muller, R. & Pennington, K., 2014. Using the Myers-Briggs Type Indicator ® and Ignatian Pedagogy Model for Improved Learning in Master's Level Finance Classes for Nurses. , 3(1), pp.64–74. Available at: http://jesuithighereducation.org/index.php/jhe/article/view/65/pdf.

Mullins-Sweatt, S.N. & Lengel, G.J., 2012. Clinical utility of the Five-Factor Model of personality disorder. *Journal of personality*, 80(6), pp.1615–39. Available at: http://www.ncbi.nlm.nih.gov/pubmed/22321379 [Accessed December 23, 2014].

Mund, M. & Neyer, F., 2014. Regulatory Focus in the Life Story: Prevention and Promotion as Expressed in Three Layers of Personality. *Journal of personality and social psychology*, 106(1), pp.169–81. Available at: http://www.ncbi.nlm.nih.gov/pubmed/24377362 [Accessed November 16, 2014].

Myers, I. et al., 1998. MBTI manual. *… and use of the Myers-Briggs Type …*. Available at: https://scholar.google.com.au/scholar?q=mbti+manual&hl=en&as_sdt=0,5#1 [Accessed March 25, 2015].

Nederhof, A., 1985. Methods of coping with social desirability bias: A review. *European Journal of Social Psychology*, 15, pp.263–280. Available at: http://onlinelibrary.wiley.com/doi/10.1002/ejsp.2420150303/abstract [Accessed October 12, 2014].

Neeley, S.M. & Cronly, M.L., 2004. When Research Participants Don't Tell It Like It Is: Pinpointing the Effects of Social Desiraibility Bias Using Self vs. Indirect-Questioning. *Advances in Consumer Research*, 31.

Nel, J.A. et al., 2012. Exploring the personality structure in the 11 languages of South Africa. *Journal of personality*, 80(4), pp.915–48. Available at: http://www.ncbi.nlm.nih.gov/pubmed/22091948 [Accessed December 23, 2014].

Nezlek, J.B. et al., 2011. A cross-cultural study of relationships between daily social interaction and the five-factor model of personality. *Journal of personality*, 79(4), pp.811–40. Available at: http://www.ncbi.nlm.nih.gov/pubmed/21682730 [Accessed January 11, 2015].

Nintendo, 2009. Wii Sports Resort. Available at: https://www.nintendo.com/games/detail/iGnKcC3xIs0WX4L3v6TedMHsQhZCKMsQ.

Ohmgane3sha, 2011. Ohm's 16-bit Minecraft Computer. Available at: https://www.youtube.com/watch?v=KzrFzkb3A4o [Accessed March 5, 2015].

Ones, D.S., Viswesvaran, C. & Reiss, A.D., 1996. Role of Social Desirability in Personality Testing for Personnel Selection: The Red Herring.pdf. *Journal of Applied Psychology*, 81(6), pp.660–679.

Park, G., Schwartz, H. & Eichstaedt, J., 2014. Automatic Personality Assessment Through Social Media Language. *Journal of personality and social psychology*. Available at: http://psycnet.apa.org/psycinfo/2014-45458-001/ [Accessed November 16, 2014].

Parmač Kovačić, M., Galić, Z. & Jerneić, Ž., 2014. Social desirability scales as indicators of self-enhancement and impression management. *Journal of personality assessment*, 96(5), pp.532–

43. Available at: http://www.ncbi.nlm.nih.gov/pubmed/24837586 [Accessed November 15, 2014].

Personality-testing.info, 2015. Personality Info - Big 5. Available at: http://personality-testing.info/tests/BIG5.php.

Pittenger, D., 2005. Measuring the MBTI and coming up short. *Consulting Psychology Journal: Practice and Research*, 57, pp.210–221. Available at: http://www.indiana.edu/~jobtalk/Articles/develop/mbti.pdf.

Play, G., 2016. Spin or Die. Available at: http://www.gojiplay.com.au/portfolio/g/.

Pope, K.S., Butcher, J.N. & Seelen, J., 2006. *The MMPI, MMPI-2, & MMPI-A in court: A practical guide for expert witnesses and attorneys* 3rd ed., Available at: http://psycnet.apa.org/psycinfo/2006-05222-000/.

Provelengios, P. & Fesakis, G., 2011. Educational applications of Serious Games : The case of the game " Food Force " in primary education students. In *5th European Conference on Games Based Learning*. pp. 476–485.

Quilty, L.C. et al., 2014. Extraversion and behavioral activation: integrating the components of approach. *Journal of personality assessment*, 96(1), pp.87–94. Available at: http://www.ncbi.nlm.nih.gov/pubmed/24063371 [Accessed November 22, 2014].

Rentfrow, P.J., Goldberg, L.R. & Levitin, D.J., 2011. The structure of musical preferences: A five-factor model. *Journal of Personality and Social Psychology*, 100(6), pp.1139–1157. Available at: /pmc/articles/PMC3138530/?report=abstract [Accessed March 15, 2015].

Rentfrow, P.J., Goldberg, L.R. & Zilca, R., 2011. Listening, watching, and reading: the structure and correlates of entertainment preferences. *Journal of personality*, 79(2), pp.223–58. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2964424&tool=pmcentrez&rendertype=abstract [Accessed February 25, 2015].

Robertson, A., 2012. Finding the Meaning in Video Games: Yes, They Have Value Beyond Entertainment and Self-Improvement. *TED Blog*. Available at: https://blog.ted.com/finding-the-meaning-in-video-games-yes-they-have-value-beyond-entertainment-and-self-improvement/.

Rooney, P., 2012. A Theoretical Framework for Serious Game Design. *International Journal of Game-Based Learning*, 2(4), pp.41–60. Available at: http://services.igi-global.com/resolvedoi/resolve.aspx?doi=10.4018/ijgbl.2012100103.

Salmond, M., 2016. *Video Game Design: Principles and Practices from the Ground Up*,

Sandal, G.M. et al., 2005. Social desirability bias in personality testing: implications for astronaut selection. *Acta astronautica*, 57(2–8), pp.634–41. Available at: http://www.ncbi.nlm.nih.gov/pubmed/16010763 [Accessed February 23, 2015].

Saucier, G. et al., 2013. A Basic Bivariate Structure of Personality Attributes Evident Across Nine Languages. *Journal of Personality*, pp.1–14. Available at: http://www.ncbi.nlm.nih.gov/pubmed/23301793 [Accessed November 13, 2014].

Seaborn, K. & Fels, D.I., 2015. Gamification in theory and action: A survey. *International Journal of Human Computer Studies*, 74, pp.14–31.

Sellbom, M., 2014. A factor mixture model approach to elaborating on offender mental health classification with the MMPI-2-RF. *Journal of personality assessment*, 96(3), pp.293–305. Available at: http://www.ncbi.nlm.nih.gov/pubmed/24148013 [Accessed November 22, 2014].

Shen, J., Brdiczka, O. & Ducheneaut, N., 2012. Inferring personality of online gamers by fusing multiple-view predictions. *User Modeling, Adaptation, and Personalization*, pp.261–273. Available at: http://link.springer.com/chapter/10.1007/978-3-642-31454-4_22 [Accessed September 17, 2014].

Sheng-Yi, H., Yu-Han, H. & Chuen-Tsai, S., 2012. Main(s) and Alts: Multiple Character

Management in World of Warcraft. In *DiGRA Nordic '12: Proceedings of 2012 International DiGRA Nordic Conference*. Available at: http://www.digra.org/digital-library/publications/mains-and-alts-multiple-character-management-in-world-of-warcraft/ [Accessed January 19, 2016].

Sipps, G.J., Alexander, R.A. & Friedt, L., 1985. Item Analysis of the Myers-Briggs Type Indicator. *Educational and Psychological Measurement*, 45(4), pp.789–796. Available at: http://epm.sagepub.com/content/45/4/789.abstract.

Software, Q.S., 2014. International personality item pool. , 8(1), pp.1–4. Available at: http://ipip.ori.org/index.htm.

Sony, 2015. PlayStation®3. *Sony Australia*. Available at: http://www.sony.com.au/productcategory/gam-ps3 [Accessed March 16, 2015].

Soto, C.J. & John, O.P., 2014. Traits in transition: the structure of parent-reported personality traits from early childhood to early adulthood. *Journal of personality*, 82(3), pp.182–99. Available at: http://www.ncbi.nlm.nih.gov/pubmed/23734942 [Accessed November 13, 2014].

Spronck, P., Balemans, I. & Van Lankveld, G., 2012. Player Profiling with Fallout 3. *Proceedings, The Eighth AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment*, pp.179–184.

Stege, L., Van Lankveld, G. & Spronck, P., 2012. Teaching high school physics with a serious game. *International Journal of Computer Science in Sport*, 11(1), pp.123–134.

Stromberg, J., 2014. Why the Myers-Briggs test is totally meaningless. *Vox Media*. Available at: http://www.vox.com/2014/7/15/5881947/myers-briggs-personality-test-meaningless [Accessed February 21, 2015].

Susi, T., Johannesson, M. & Backlund, P., 2007. Serious Games – An Overview.

Tackett, J.L. et al., 2012. The hierarchical structure of childhood personality in five countries: continuity from early childhood to early adolescence. *Journal of personality*, 80(4), pp.847–79.

Available at: http://www.ncbi.nlm.nih.gov/pubmed/22091829 [Accessed December 23, 2014].

Tekofsky, S. et al., 2015. Past Our Prime: A Study of Age and Play Style Development in Battlefield 3. *IEEE Transactions on Computational Intelligence and AI in Games*, 7(3), pp.292–303. Available at: http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7012062&isnumber=7254210.

Tekofsky, S., Spronck, P., Plaat, A., Van Den Herik, J., et al., 2013. Play style: Showing your age. *IEEE Conference on Computatonal Intelligence and Games, CIG*.

Tekofsky, S., Spronck, P., Plaat, A., Herik, J. van den, et al., 2013. PsyOps : Personality Assessment Through Gaming Behavior. *International Conference on the Foundations of Digital Games*, (June 2014), pp.166–173.

Tellegen, A. & Atkinson, G., 1974. Openness to absorbing and self-altering experiences ("absorption"), a trait related to hypnotic susceptibility. *Journal of Abnormal Psychology*, 83(3), pp.268–277. Available at: http://psycnet.apa.org/psycinfo/1974-32492-001.

Trull, T.J., 2012. The Five-Factor Model of personality disorder and DSM-5. *Journal of personality*, 80(6), pp.1697–720. Available at: http://www.ncbi.nlm.nih.gov/pubmed/22321203 [Accessed December 23, 2014].

Tupes, E.C. & Christal, R.E., 1961. *Recurrent Personality Factors Based on Trait Ratings*,

UNWF, 2005. Food Force. *United Nations World Food Programme*. Available at: http://www.download-free-games.com/freeware_games/food_force.htm [Accessed November 3, 2017].

Vazire, S., 2010. Who knows what about a person? The self–other knowledge asymmetry (SOKA) model. *Journal of personality and social psychology*, 98(2), pp.281–300. Available at: http://psycnet.apa.org/journals/psp/98/2/281/.

Widiger, T. a & Costa, P.T.J., 2012. Integrating normal and abnormal personality structure: the Five-Factor Model. *Journal of personality*, 80(6), pp.1471–506. Available at:

http://www.ncbi.nlm.nih.gov/pubmed/22320149 [Accessed December 23, 2014].

Widiger, T. a & Presnall, J.R., 2013. Clinical application of the five-factor model. *Journal of personality*, 81(6), pp.515–27. Available at: http://www.ncbi.nlm.nih.gov/pubmed/22924994 [Accessed December 9, 2014].

Woo, S.E. et al., 2014. Validity of six openness facets in predicting work behaviors: a meta-analysis. *Journal of personality assessment*, 96(1), pp.76–86. Available at: http://www.ncbi.nlm.nih.gov/pubmed/23795997 [Accessed November 13, 2014].

Yang, G.S. et al., 2014. Effects of Avatar Race in Violent Video Games on Racial Attitudes and Aggression. *Social Psychological and Personality Science*, 5(6), pp.698–704. Available at: http://spp.sagepub.com/cgi/doi/10.1177/1948550614528008 [Accessed November 17, 2014].

Yang, K. & Bond, M.H., 1990. Exploring implicit personality theories with indigenous or imported constructs: the Chinese case. *Journal of personality and social psychology1*, 58(6), pp.1087–1095. Available at: http://psycnet.apa.org/journals/psp/58/6/1087/.

York, D.J., 2014. A Review of Fundamentals of MMPI–2 Interpretation: The Eight Basic Clinical Scales. *Journal of Personality Assessment*, 96(4), pp.478–479. Available at: http://www.tandfonline.com/doi/abs/10.1080/00223891.2014.896370 [Accessed November 22, 2014].

Zyda, M., 2005. From visual simulation to virtual reality to games. *Computer*, 38(9), pp.25–32. Available at: http://ieeexplore.ieee.org.ezproxy.ugm.ac.id/ielx5/2/32339/01510565.pdf?tp=&arnumber=1510565&isnumber=32339%5Cnhttp://ieeexplore.ieee.org.ezproxy.ugm.ac.id/xpl/articleDetails.jsp?tp=&arnumber=1510565&queryText%3DSimulation+to+Virtual+Reality+to+Games.

8. Appendix

8.1 Experiment Demographics

ID	Age	Gender	Experience With Games	Average Time Playing Games	Highest Education Level
26	43	Male	> 10 years	4-6 hours a week	University
216	31	Male	> 10 years	0-2 hours a week	University
417	44	Male	> 10 years	6-8 hours a week	University
545	25	Male	> 10 years	8+ hours a week	University
564	19	Male	> 10 years	8+ hours a week	University
658	22	Male	> 10 years	8+ hours a week	University
1016	20	Male	> 10 years	8+ hours a week	University
1097	22	Male	5-10 years	8+ hours a week	College
1219	23	Male	> 10 years	8+ hours a week	University
1526	31	Male	> 10 years	4-6 hours a week	University
2186	30	Female	> 10 years	8+ hours a week	University
2452	34	Male	> 10 years	8+ hours a week	University
2675	29	Female	> 10 years	8+ hours a week	Postgraduate
2688	23	Male	> 10 years	8+ hours a week	High School
3184	25	Male	> 10 years	8+ hours a week	University
3423	30	Male	> 10 years	2-4 hours a week	Vocational College
3482	60	Female	< 1 year	0-2 hours a week	University
4070	33	Male	> 10 years	8+ hours a week	University
4102	29	Male	> 10 years	6-8 hours a week	University
4406	39	Female	> 10 years	6-8 hours a week	University
4507	39	Male	> 10 years	4-6 hours a week	University
4717	38	Male	> 10 years	8+ hours a week	University
5053	23	Other	> 10 years	2-4 hours a week	Postgraduate
5762	27	Male	> 10 years	8+ hours a week	University
6071	25	Male	> 10 years	8+ hours a week	University
6235	16	Male	5-10 years	2-4 hours a week	High School
6261	24	Male	> 10 years	8+ hours a week	University
6454	23	Female	2-5 years	0-2 hours a week	Postgraduate
6763	23	Male	> 10 years	2-4 hours a week	University
6930	26	Male	> 10 years	8+ hours a week	University
7383	29	Male	> 10 years	2-4 hours a week	University
7630	60	Female	2-5 years	0-2 hours a week	University
7731	33	Male	> 10 years	4-6 hours a week	Postgraduate
7951	37	Male	> 10 years	4-6 hours a week	College
8330	36	Male	> 10 years	0-2 hours a week	Postgraduate
8339	18	Male	2-5 years	4-6 hours a week	High School
8395	22	Male	> 10 years	8+ hours a week	College

8486	21	Male	> 10 years	2-4 hours a week	University
8668	27	Male	> 10 years	4-6 hours a week	University
9048	25	Male	> 10 years	4-6 hours a week	University
9486	21	Male	> 10 years	8+ hours a week	University
9855	29	Male	> 10 years	6-8 hours a week	University
9878	32	Male	> 10 years	0-2 hours a week	University

8.2 IPIP NEO-PI-R Scoring Keys

8.2.1 Neuroticism

N1: Anxiety	**+ keyed**	**Worry about things.**
N1: Anxiety	+ keyed	Fear for the worst.
N1: Anxiety	+ keyed	Am afraid of many things.
N1: Anxiety	+ keyed	Get stressed out easily.
N2: Anger	+ keyed	Get angry easily.
N2: Anger	+ keyed	Get irritated easily.
N2: Anger	+ keyed	Lose my temper.
N2: Anger	– keyed	Am not easily annoyed.
N3: Depression	+ keyed	Often feel blue.
N3: Depression	+ keyed	Dislike myself.
N3: Depression	+ keyed	Am often down in the dumps.
N3: Depression	– keyed	Feel comfortable with myself.
N4:Self-Consciousness	+ keyed	Find it difficult to approach others.
N4:Self-Consciousness	+ keyed	Am afraid to draw attention to myself.
N4:Self-Consciousness	+ keyed	Only feel comfortable with friends.
N4:Self-Consciousness	– keyed	Am not bothered by difficult social situations.
N5: Immoderation	+ keyed	Go on binges.
N5: Immoderation	– keyed	Rarely overindulge.
N5: Immoderation	– keyed	Easily resist temptations.
N5: Immoderation	– keyed	Am able to control my cravings.
N6: Vulnerability	+ keyed	Panic easily.
N6: Vulnerability	+ keyed	Become overwhelmed by events.
N6: Vulnerability	+ keyed	Feel that I'm unable to deal with things.
N6: Vulnerability	– keyed	Remain calm under pressure.

Factor of Neuroticism

8.2.2 Extraversion

E1: Friendliness	**+ keyed**	**Make friends easily.**
E1: Friendliness	+ keyed	Feel comfortable around people.
E1: Friendliness	– keyed	Avoid contacts with others.
E1: Friendliness	– keyed	Keep others at a distance.
E2: Gregariousness	+ keyed	Love large parties.
E2: Gregariousness	+ keyed	Talk to a lot of different people at parties.
E2: Gregariousness	– keyed	Prefer to be alone.
E2: Gregariousness	– keyed	Avoid crowds.
E3: Assertiveness	+ keyed	Take charge.
E3: Assertiveness	+ keyed	Try to lead others.
E3: Assertiveness	+ keyed	Take control of things.
E3: Assertiveness	– keyed	Wait for others to lead the way.
E4: Activity Level	+ keyed	Am always busy.
E4: Activity Level	+ keyed	Am always on the go.
E4: Activity Level	+ keyed	Do a lot in my spare time.
E4: Activity Level	– keyed	Like to take it easy.
E5: Excitement-Seeking	+ keyed	Love excitement.
E5: Excitement-Seeking	+ keyed	Seek adventure.
E5: Excitement-Seeking	+ keyed	Enjoy being reckless.
E5: Excitement-Seeking	+ keyed	Act wild and crazy.
E6:Cheerfulness	+ keyed	Radiate joy.
E6:Cheerfulness	+ keyed	Have a lot of fun.
E6:Cheerfulness	+ keyed	Love life.
E6:Cheerfulness	+ keyed	Look at the bright side of life.

Factor of Extraversion

8.2.3 Openness

O1: Imagination	+ keyed	Have a vivid imagination.
O1: Imagination	+ keyed	Enjoy wild flights of fantasy.
O1: Imagination	+ keyed	Love to daydream.
O1: Imagination	+ keyed	Like to get lost in thought.
O2: Artistic Interests	+ keyed	Believe in the importance of art.
O2: Artistic Interests	+ keyed	See beauty in things that others might not notice.
O2: Artistic Interests	– keyed	Do not like poetry.
O2: Artistic Interests	– keyed	Do not enjoy going to art museums.
O3: Emotionality	+ keyed	Experience my emotions intensely.
O3: Emotionality	+ keyed	Feel others' emotions.
O3: Emotionality	– keyed	Rarely notice my emotional reactions.
O3: Emotionality	– keyed	Don't understand people who get emotional.
O4: Adventurousness	+ keyed	Prefer variety to routine.
O4: Adventurousness	– keyed	Prefer to stick with things that I know.
O4: Adventurousness	– keyed	Dislike changes.
O4: Adventurousness	– keyed	Am attached to conventional ways.
O5: Intellect	+ keyed	Love to read challenging material.
O5: Intellect	– keyed	Avoid philosophical discussions.
O5: Intellect	– keyed	Have difficulty understanding abstract ideas.
O5: Intellect	– keyed	Am not interested in theoretical discussions.
O6: Liberalism	+ keyed	Tend to vote for liberal political candidates.
O6: Liberalism	+ keyed	Believe that there is no absolute right and wrong.
O6: Liberalism	– keyed	Tend to vote for conservative political candidates.
O6: Liberalism	– keyed	Believe that we should be tough on crime.

Factor of Openness to Experience

8.2.4 Agreeableness

A1: Trust	+ keyed	Trust others.
A1: Trust	+ keyed	Believe that others have good intentions.
A1: Trust	+ keyed	Trust what people say.
A1: Trust	– keyed	Distrust people.
A2: Morality	– keyed	Use others for my own ends.
A2: Morality	– keyed	Cheat to get ahead.
A2: Morality	– keyed	Take advantage of others.
A2: Morality	– keyed	Obstruct others' plans.
A3: Altruism	+ keyed	Am concerned about others.
A3: Altruism	+ keyed	Love to help others.
A3: Altruism	– keyed	Am indifferent to the feelings of others.
A3: Altruism	– keyed	Take no time for others.
A4: Cooperation	– keyed	Love a good fight.
A4: Cooperation	– keyed	Yell at people.
A4: Cooperation	– keyed	Insult people.
A4: Cooperation	– keyed	Get back at others.
A5: Modesty	– keyed	Believe that I am better than others.
A5: Modesty	– keyed	Think highly of myself.
A5: Modesty	– keyed	Have a high opinion of myself.
A5: Modesty	– keyed	Boast about my virtues.
A6: Sympathy	+ keyed	Sympathize with the homeless.
A6: Sympathy	+ keyed	Feel sympathy for those who are worse off than myself.
A6: Sympathy	– keyed	Am not interested in other people's problems.
A6: Sympathy	– keyed	Try not to think about the needy.

Factor of Agreeableness

8.2.5 Conscientiousness

C1: Self-Efficacy	+ keyed	Complete tasks successfully.
C1: Self-Efficacy	+ keyed	Excel in what I do.
C1: Self-Efficacy	+ keyed	Handle tasks smoothly.
C1: Self-Efficacy	+ keyed	Know how to get things done.
C2: Orderliness	+ keyed	Like to tidy up.
C2: Orderliness	– keyed	Often forget to put things back in their proper place.
C2: Orderliness	– keyed	Leave a mess in my room.
C2: Orderliness	– keyed	Leave my belongings around.
C3: Dutifulness	+ keyed	Keep my promises.
C3: Dutifulness	+ keyed	Tell the truth.
C3: Dutifulness	– keyed	Break rules.
C3: Dutifulness	– keyed	Break my promises.
C4: Achievement-Striving	+ keyed	Do more than what's expected of me.
C4: Achievement-Striving	+ keyed	Work hard.
C4: Achievement-Striving	– keyed	Put little time and effort into my work.
C4: Achievement-Striving	– keyed	Do just enough work to get by.
C5: Self-Discipline	+ keyed	Am always prepared.
C5: Self-Discipline	+ keyed	Carry out my plans.
C5: Self-Discipline	– keyed	Waste my time.
C5: Self-Discipline	– keyed	Have difficulty starting tasks.
C6: Cautiousness	– keyed	Jump into things without thinking.
C6: Cautiousness	– keyed	Make rash decisions.
C6: Cautiousness	– keyed	Rush into things.
C6: Cautiousness	– keyed	Act without thinking.

Factor of Conscientiousness

8.3 Calculated Facet Scores From Experiment

ID	N1	N2	N3	N4	N5	N6	E1	E2	E3	E4	E5	E6	O1	O2	O3	O4	O5	O6	A1	A2	A3	A4	A5	A6	C1	C2	C3	C4	C5	C6
26	13	11	9	15	13	12	12	11	15	12	12	16	11	17	12	11	17	16	14	12	14	15	9	14	16	13	17	14	12	14
216	9	7	10	10	18	9	19	18	18	12	17	19	18	15	20	18	19	15	20	18	16	12	8	17	19	7	19	12	10	8
417	18	12	12	17	11	13	8	9	11	15	11	13	20	14	16	12	17	14	8	17	15	15	11	16	15	13	15	16	11	17
545	6	5	5	6	13	5	16	15	13	10	14	15	16	8	8	13	14	11	6	16	12	13	14	9	13	6	16	11	11	8
564	18	11	11	8	12	15	16	12	11	13	11	20	20	16	17	14	15	10	12	14	18	16	10	14	12	10	13	18	10	10
658	14	11	16	20	14	10	10	8	13	13	12	14	13	16	14	11	16	11	11	17	18	14	11	12	12	17	17	14	6	15
1016	20	9	18	16	18	18	14	10	9	11	14	9	19	9	13	7	17	6	14	17	16	14	9	14	11	7	15	7	9	9
1097	9	6	8	14	13	5	12	7	11	11	13	17	20	10	8	14	19	13	13	13	9	16	12	11	17	7	18	9	11	12
1219	11	15	10	11	12	8	13	9	10	12	13	14	13	8	14	11	14	11	11	12	14	8	12	11	16	12	13	15	11	13
1526	19	11	18	13	14	14	11	11	10	10	9	9	10	14	15	9	15	13	12	16	18	16	19	15	13	12	18	15	12	14
2186	15	12	12	15	14	9	9	10	13	14	12	14	20	18	15	14	17	13	12	17	18	11	19	18	15	12	15	17	13	13
2452	15	14	15	18	12	9	5	5	12	13	11	10	16	15	9	17	19	17	5	13	14	11	16	15	13	7	11	17	12	15
2675	15	8	8	17	17	10	12	5	8	14	12	15	15	15	16	9	16	19	14	18	16	18	16	17	16	8	16	16	12	16
2688	5	13	11	14	10	7	10	7	17	8	16	15	13	6	13	10	9	12	9	12	12	10	8	7	18	18	19	9	8	17
3184	17	9	11	12	12	9	13	13	16	15	14	15	18	16	14	14	19	15	12	18	17	17	16	14	15	10	15	14	14	15
3423	17	6	16	20	13	14	8	9	7	16	6	8	16	13	17	18	11	18	14	18	18	19	20	14	13	6	17	17	10	18
3482	10	7	4	11	11	11	19	11	11	12	9	18	9	11	15	10	11	12	17	19	19	18	17	13	13	16	20	14	14	15
4070	18	13	16	18	19	17	6	5	11	10	6	7	13	9	13	6	10	14	8	17	15	18	13	10	14	9	15	14	8	15
4102	11	16	13	15	20	12	11	8	13	9	13	12	19	15	14	14	18	15	16	12	14	15	9	14	10	12	15	8	8	9
4406	20	16	15	19	16	12	9	4	15	12	10	10	20	16	18	13	20	11	13	19	15	17	14	11	20	12	18	19	13	9
4507	10	13	9	6	16	6	20	15	18	16	14	19	16	18	12	18	19	18	20	19	19	20	10	16	16	8	17	20	14	11
4717	9	9	15	4	18	5	20	17	19	10	18	14	19	17	18	18	20	15	14	20	20	14	20	20	12	12	16	15	10	10
5053	13	11	12	10	12	8	15	11	16	14	15	13	15	17	16	16	15	15	13	16	17	15	11	17	17	14	14	16	15	10
5762	20	16	20	19	12	19	4	12	11	7	12	4	16	13	9	9	15	9	5	11	17	9	11	15	11	11	14	9	10	15
6071	6	9	8	9	13	6	14	6	18	8	11	15	20	19	9	9	19	14	10	14	14	12	12	17	17	8	15	15	11	15
6235	12	8	10	11	13	13	14	14	5	7	16	14	17	11	13	5	10	11	15	16	17	15	14	10	14	9	13	10	10	16
6261	9	13	7	12	11	13	12	9	9	12	13	15	17	10	9	12	12	10	14	17	13	12	14	14	11	17	16	9	10	12
6454	15	12	8	11	10	9	15	13	14	14	15	19	16	15	15	12	14	12	18	20	20	18	18	18	17	17	19	18	16	16
6763	15	19	6	18	18	16	10	9	10	9	15	15	20	20	17	8	20	16	11	14	19	8	9	15	20	9	17	8	12	4
6930	13	4	12	16	12	12	15	11	7	6	14	13	17	15	17	11	15	14	16	18	18	16	17	17	15	11	16	13	10	11
7383	16	11	16	16	13	8	12	10	12	11	9	8	20	17	19	11	18	12	14	16	20	19	17	17	13	6	19	14	12	16
7630	10	7	5	12	11	9	14	9	11	16	8	16	8	12	14	13	13	8	14	20	15	20	16	15	15	15	18	15	14	15
7731	11	9	10	16	9	9	8	11	7	11	11	13	19	19	14	14	20	15	15	17	18	16	13	16	15	18	16	18	9	20
7951	17	6	15	15	20	14	9	9	9	14	20	16	20	13	16	16	20	15	12	14	19	19	18	20	14	20	8	18	10	5
8330	11	7	8	12	10	10	15	11	16	16	11	18	16	16	15	13	18	16	16	18	16	18	13	15	17	11	15	18	14	11
8339	13	13	12	12	12	14	12	12	14	11	13	12	12	13	13	11	11	10	15	18	15	17	16	15	13	11	14	12	13	15
8395	15	10	13	16	10	13	15	10	11	11	17	16	16	19	14	8	13	9	17	12	18	14	20	17	16	13	16	12	10	9
8486	13	8	15	17	8	7	9	14	15	12	11	12	12	8	14	18	12	11	7	20	18	18	19	15	15	19	19	18	16	16
8668	8	13	7	13	17	9	13	5	10	10	6	11	16	16	13	13	16	18	14	12	17	14	7	12	16	6	13	14	12	17
9048	14	13	5	9	9	9	17	8	16	14	15	18	16	17	11	13	16	13	15	18	15	14	8	12	16	14	16	14	11	11
9486	9	8	6	11	10	7	13	10	11	9	13	14	16	12	13	11	17	18	14	14	13	12	12	7	14	16	16	10	9	14
9855	11	9	17	19	18	8	10	11	16	12	15	14	19	15	17	13	20	19	4	8	18	16	10	15	17	17	11	15	11	13
9878	12	12	5	9	11	7	17	11	13	15	8	12	17	13	17	9	14	10	14	20	17	17	12	15	17	19	15	19	14	16

8.4 Calculated Factor Scores From Experiment

Player ID	Neuroticism	Extraversion	Openness	Agreeableness	Conscientiousness
26	73	78	84	78	86
216	63	103	105	91	75
417	83	67	93	82	87
545	40	83	70	70	65
564	75	83	92	84	73
658	85	70	81	83	81
1016	99	67	71	84	58
1097	55	71	84	74	74
1219	67	71	71	68	80
1526	89	60	76	96	84
2186	77	72	97	95	85
2452	83	56	93	74	75
2675	75	66	90	99	84
2688	60	73	63	58	89
3184	70	86	96	94	83
3423	86	54	93	103	81
3482	54	80	68	103	92
4070	101	45	65	81	75
4102	87	66	95	80	62
4406	98	60	98	89	91
4507	60	102	101	104	86
4717	60	98	107	108	75
5053	66	84	94	89	86
5762	106	50	71	68	70
6071	51	72	90	79	81
6235	67	70	67	87	72
6261	65	70	70	84	75
6454	65	90	84	112	103
6763	92	68	101	76	70
6930	69	66	89	102	76
7383	80	62	97	103	80
7630	54	74	68	100	92
7731	64	61	101	95	96
7951	87	77	100	102	75
8330	58	87	94	96	86
8339	76	74	70	96	78
8395	77	80	79	98	76
8486	68	73	75	97	103
8668	67	55	92	76	78
9048	59	88	86	82	82
9486	51	70	87	72	79
9855	82	78	103	71	84
9878	56	76	80	95	100