EVALUATING THE USER-EXPERIENCE OF EXISTING STRATEGIES TO LIMIT VIDEO GAME SESSION LENGTH

Bryan Davies

Department of Computer Science

University of Cape Town

This dissertation is submitted for the degree of Master of Science

March 2018

The financial assistance of the National Research Foundation (NRF) towards this research is hereby acknowledged. Opinions expressed and conclusions arrived at, are those of the author and are not necessarily to be attributed to the NRF

To my parents Liz and Neil Davies,

For their unwavering support,

For the sacrifices they’ve made,

For making me, *me.*

ABSTRACT

Digital video games are an immensely popular form of entertainment. The meaningful positive experiences that games facilitate are fundamental to the activity; players are known to invest a lot of time playing games in search of those experiences. Digital games research is polarized. Some studies find games to be a healthy hobby with positive effects; games promote well-being through regular experience of positive psychological experiences such as flow and positive emotions. Others have identified rare problematic use in those players who devote excessive amounts of time to gaming, associating them with social dysfunction, addiction, and maladaptive aggression. While it remains unclear if games cause these effects, or merely coincide with play, the negative effects historically receive more attention in both popular media and academia.

Some authorities attempt to reduce the harms associated with games to such an extent that their methods have become national policy affecting all players including those who exhibit no negative outcomes. In South Korea and Taiwan, policing authorities employ a behaviour policy that sets strict daily limits on session length, thereby controlling the amount of time people spend playing games each day. In China, the General Administration of Press and Publication employ a design policy requiring that games service-providers fatigue their games' mechanics after a period to coax them to take a break sooner than they ordinarily would. Both policy types alter player interaction with games in any given session and it is unclear how these policies affect players in general.

This research aims to compare sessions affected by the behaviour policy, design policy, and policy-free sessions in terms of session length, measurable subjective user-experience, the player's intention to return to the game, and their reasons for choosing to stop playing in a particular session. For use in a repeated-measures experiment, we modified the action RPG Torchlight II to simulate both policies. Participants had one session at the same time each week for three consecutive weeks. In varied sequences, participants played a control session unaffected by policy, a one-hour shutdown session representing behaviour policy, and a fatigue session representing design policy. After each session, we recorded their session's length, their user-experience in terms of flow and affect, their intention to return to the game, and their reason for ending the current session.

We found that our shutdown condition successfully decreases session length, when compared to the other conditions. The condition facilitates strong flow, moderate positive-affect, and weak negative-affect. The shutdown event does not appear to degrade positive experiences and makes participants slightly more upset (statistically significant) than they would be after choosing to stop playing. This is because players do not get to make that decision, and because players are unable to complete the goals they have set for themselves. Most players intended to play the game again immediately or sometime later in that same day, much sooner (statistically significant)

than they would after choosing to stop. This also may be due to satisfaction associated with choosing to stop, or being unable to complete their self-set goals.

We found that our fatigue condition increases session length when compared to the other conditions. This result contradicts the intentions associated with design policies: shorter sessions. The fatigue mechanics make the game more difficult, which increases the time required for players to complete the goals they have set for themselves, whether it is to complete a level, quest, or narrative sequence. The condition facilitates high levels of flow, moderate positive-affect and low negative-affect; the condition does not appear to degrade these positive experiences, nor increase negative experience. Most players intended to take the longest breaks between sessions of at least one day, and although we observed that these were longer than the control condition, the differences is not statistically significant.

We found that most participants chose to stop playing when the game stopped providing them with positive experiences, or begins to generate discomfort. A large group of participants chose to stop because another activity took priority. Few participants chose to stop because they were satisfied with their session. Less than one third of players explicitly referenced the fatigue mechanics in their decision to stop. Neither policy is holistically better than the other. Both provide strong positive experiences, and have different effects on session length. Whereas it appears that the fatigue condition fails to reduce session length, it also appears that players intend taking longer breaks between sessions, which may reduce total play-time across all sessions. Similarly, the shutdown condition may increase total play time, or at least bring it closer to normal amounts of play-time while also making players more upset.

Our operational definition of user-experience is bi-dimensional, and does not include many experiential constructs commonly associated with digital games. During this research, several reliable and valid, and more representative experience measures became available. Any future work on this topic should make use of one of these. Our experiment tested the effects of player experience associated with a single game, genre, and context. Future research should reduce the variation of player factors by focusing on single personalities, typologies, or risk-factors rather than generalizing to all players. We tested out participants only as they played in the early stages of Torchlight II. It is possible that the game's narrative elements, rather than the gameplay mechanics fatigued by the design policy, motivated continued play. We suggest a longitudinal study of the individual policies to explore their effects over many sessions.

ACKNOWLEDGEMENTS

This project has not always been good to me. There were times when it seemed most sensible to quit chasing truth and lapse into ignorance, chasing bliss by accepting mystery. But I did not slip, primarily because people were always good to me.

My greatest thanks go to my supervisor, Edwin Blake, for allowing me to pursue this somewhat out-of-place pet-subject, and then guiding me to a manageable solution for every new problem I invented. I will never forget his continued understanding and invaluable support throughout these years, in every achievement and set-back, academic and otherwise.

To those starring in their own post-graduate productions, who still found time to wait in the wings of my play: Josh, Ernesto, Anja, Ryan, Jacob, and Marcel, I thank you for cheering me on, suggesting new interpretations, and prompting me whenever I botched a reading. I thank you for giving me inspiration, affirmation, recommendation and kicks-in-the-shin; I needed every single one.

To my friends and family, my willing practice audience at every stage of this project. I wish I could repay for each read and each listen, for every meal and every much-needed distraction in the wide world outside the lab. You may not have been directly involved in this project, but you invested in me, and I cannot thank you enough for giving me so much of you.

CONTENTS

List of Tables

List of Figures

LIST OF ABBREVIATIONS AND ACRONYMS

UXG: User-experience in digital games

MMO: Massively-multiplayer online (game)

RPG: Role-playing game

FSS: Flow Short Scale

PA: Positive-affect

NA: Negative-affect

I-PANAS-SF: International Positive And Negative Affect Schedule Short Form

RELEVANT GLOSSARY OF TERMS AND CONCEPTS

Well-being: Satisfaction with quality of life, health, or happiness.

Game: A goal-oriented, rule-bound activity that one engages with for amusement.

Play: The act of performing a game's activity or activities.

Player: The agent of play.

Entertainment game: A game designed to provide intrinsic entertainment through play. We exclude gambling games, which provide extrinsic rewards.

Digital games: An entertainment game played in digital format including, but not limited to, personal computers, game consoles and mobile devices.

Ludology: The study of games and play

Virtual Environment (VE): Digital interactive space; the space in which digital games manifest. Represented graphically in two- or three-dimensional digital space.

Virtual World: An expansive virtual environment that simulates the real world at varying fidelity.

Avatar: A virtual representation of a player within a Virtual Environment

Single-player game: A game that responds to one active player in the virtual world at any given time.

Multi-player game: A game that responds to multiple active players at any given time.

Online game: A game that necessarily requires an active internet connection for play.

Offline game: A game that does not necessarily require an active internet connection for play. Many offline games have online game modes.

Game Mechanic: Sets of rules that govern how the player interacts within a virtual environment allowing gameplay.

Genre: The classification of digital games. A game fits into a genre based on the mechanics, rules, and challenges that it presents.

Action Game: Games that require players perform physical actions and reactions from the player to perform the game mechanics. Distinct from games that only require the player to navigate a menu.

Shooter: Games that have shooting as a core mechanic. Players must use a projectile tool, often weapons, to overcome challenges. Shooters often simulate firearm combat.

Role-Playing Game (RPG): Games that allows players to act as an agent in a game world; Players make decisions that affect the character in terms of body, ability, and personality.

Session: the period during which playing a game is the player's primary focus.

Session length: the length, in time, given to a gameplay session

Hiatus: an indefinite pause or break in an activity, such as gameplay

Hiatus length: the length in time given to the hiatus.

Normal gameplay: playing a game with the designer's original mechanics

User-Experience in Digital Games (UXG): A set of psychological constructs that describe the subjective experiences associated with playing digital games

Flow: A state of mind that describes the positive experience of activities associated with focussed attention, motivation, enjoyment and time dilation. Often called the *optimal experience.*

Chronoslip: Dilation in perception of time, when time seems to move faster (Parkin, 2015).

Affect: The physiological presentation and subject psychological experience of emotions.

Experience Points (XP): A common game mechanic that is usually a measure of character progression or experience of a player associated with a game account. Players gain XP by playing the game (or games) with which the XP is associated.

Loot: all in-game artefacts that a player can use to develop their character in terms of strength or identity. Includes the currency of the virtual world, the clothes, armour and weapons that a character might use, XP, and trophies, titles and any artefact that grants renown.

Respawn: The process by which an avatar is resurrected.

Fighting: The primary activity in action RPGs. Players battle monsters to progress, and gain loot.

Fishing: An activity in the Torchlight II virtual world.

Exploring: The act of moving around the virtual world, seeing what there is to see.

Questing: Performing tasks or actions that progress the story or narrative. Often gives small amounts of loot upon completion.

List of Appendices

1 INTRODUCTION

Digital games are an immensely popular and widely celebrated entertainment medium with an estimated global player-base of 1,7 billion people (Newzoo, 2014). Games players have been known to invest a lot of time in the activity due to the fun, meaningful and compelling experiences they facilitate; games are fundamentally an experiential activity. Some people who play games, especially those who spend the most time playing, experience a variety of problems including social dysfunction (Kowert, 2016), poor health (Olson, 2016), and addiction(Griffiths, 2016). The link between games and these problems is apparent, but it remains unclear if they are the cause of these issues. In light of such problems, some authorities attempt to reduce the harms apparently caused by games to such an extent that their methods have become national policy (Lee, 2014; see Section 1.2). Some policies control the amount of time people spend playing games by setting strict daily limits on session length (see Section 1.2.1). Others attempt to influence player decisions, getting them to choose an earlier exit from the game than they ordinarily would (see Section 1.2.2). Both practices change how players interact with games to some extent, but each practice's actual effects on session length and the player's experiences remain unclear and we would like to know what these effects are.

In this chapter, we present a brief overview of the positive and negative effects of digital games. We describe two national scale policies aimed at minimizing the negative effects, and identify some problems with the policies. We briefly address the concept of user-experience in digital games before describing our research questions in detail.

1.1 Digital games: an affecting medium

Games are a diverse entertainment medium that has the potential to affect its players in both positive and negative ways. Digital games can produce compelling positive experiences that may have lasting benefits on the social, cognitive, and emotional aspects of players' lives (Granic, Lobel, & Engels, 2014). Some games are fun and amusing. Others evoke strong emotional responses. Some games maintain expansive virtual worlds in which players can immerse themselves. An increasing number of eSports games provide fiercely competitive environments for players. Most games have the potential to change the lives of those who play them. Granic and colleagues (2014) propose that games owe their popularity to their *affectivity*, and the myriad compelling and meaningful *experiences* that the medium provides. Games would be unpopular if players did not value the experiences evoked while playing. In fact, players are motivated to continue playing games because of strong positive psychological experiences that games facilitate (Ryan, Rigby, & Przybylski, 2006) including enjoyment (Bowman, Weber, Tamborini, & Sherry, 2013; Sherry, 2004; Sweetser & Wyeth, 2005), needs satisfaction such as feelings of autonomy and competence (Ryan et al., 2006; Tamborini, Bowman, Eden, Grizzard, & Organ, 2010; Nick Yee, 2006), and a range of emotions such as *fiero*[1] that are not accessible through traditional media such as film and literature (Lazzaro, 2004). See Section 2.2.1 for more discussion about the beneficial effects of digital games.

As popular as games may be, it is they are associated with health risks and behaviours sometimes considered to be immoral, such as addiction and aggression (Bowman, 2016; Granic et al., 2014). Games have a way of distracting players from noticing the passage of time, and because games are not considered to be virtuous, time spent playing is considered a "tremendous waste of time" (Parkin, 2015, p. 5). This opinion is reflected in the early literature, where authors chose to use the words "obsession"(Ross, Finestone, & Lavin, 1982), "catatonia" (Nills (1982), as cited in Griffiths, 2015), and "junkies" and "addiction" (Soper & Miller, 1983) to describe players and their relationship to video games. These reports generated concern in various responsible parties, including parents, educators and government and inspired research that investigates the idea of gaming related disorders, and continues to this day. The research is increasingly sophisticated (Kowert & Quandt, 2016), but there remain significant concerns that the problems associated with video game-play are not caused by games. Debate about the classification of gaming related problems as a new disorder, separate from existing disorders continues (Aarseth et al., 2016).

Arguably, the most heated debate revolves around violence and aggression, particularly in youths. Violent and aggressive play is no new phenomenon. Traditional games, including tag, cops and robbers, and red rover act as safe facsimiles for unsafe and potentially violent situations. Some of the most popular contemporary digital games are facsimiles of contemporary risks: war, gangsterism and gun violence (Coulson & Ferguson, 2016). Furthermore, because contemporary games of this type provide high-fidelity visual representations of violence, it is important to question their effect on players, especially youths in developmental phases.

[1] Personal triumph over adversity.

Perhaps it is the widespread publication of game related tragedies that has instilled the collective belief that digital games pose a severe and genuine threat to society as a whole. Reports link video games to matricide (Burge, 2009), child assault (Fleming, 2011), fatal exhaustion (Prynne, 2014), and mass murder Pidd, 2012; THR Staff, 2013). Elson and Ferguson (2013), and Bowman (2016) liken the current state of the discourse to a *moral panic*; a situation wherein society finds a social object (such as digital games) and blames it for immoral acts and tragedies. In doing so, society ignores the complex situations that precipitate them, such as mental health and social dynamics. Furthermore, popular science reporting has a habit of using scientific findings to confirm suspected fears, further vilify the object, and validate existing panic (Elson & Ferguson, 2013). That being said, moral panics could benefit society. Widespread awareness of an issue can influence scientific research because the questions raised by society are pertinent, and the answers are unclear – as it is with digital games. Unfortunately, behavioural research often takes time to formulate conclusions on a matter, and useful interventions require these conclusions. Either way, moral panic prompts social action taken to solve the problem even if the outcome is unknown.

1.2 Taking action and session length policies

“You’ve been playing for a while. Why not take a break?” -Nintendo

Various societies acted in response to the perceived risks associated with playing digital games. In the United States of America and European Union respectively, non-governmental ratings authorities such as the ESRB and PEGI formed systems to address adult content in digital games and their effects. Many other societies recommend or practice some variation on the theme of reducing screen-time, or *session length*. It makes sense that time spent playing games steals time from the acceptable pursuits of contemporary society, such as academia, socializing, and organized sports. Frequent allusions to the idea of game addiction prime authorities and policy makers to approach problematic gaming behaviour as they would a substance abuse problem: by promoting moderation or abstinence. The idea being that when one ceases to consume the substance then the symptoms will fade, viz. if one stops playing games, the negative effects will cease. This strict moderation practice is often used in the home environment where parents or guardians limit session length (Entertainment Software Association, 2016). South Korea, Taiwan and China notably introduced policies to moderate session length with the intention of reducing problematic game-play and bullying (Yuan & Lee, 2014).

1.2.1 Behaviour policies

Some authorities monitor and moderate player behaviour to minimize the incidence of problematic play. Nearly 80% of US parents place time limits on digital game play (Entertainment Software Association, 2016). South Korean government passed the Youth Protection Revision Bill in 2011, the so-called *shutdown* law that prohibits all persons under the age of sixteen from playing online-games between the hours of midnight and 6am. In 2012, the South Korean Ministry of Education, Science, and Technology announced the *cooling off* law that limits how much time gamers could spend playing in any one 24-hour period: The game forces players to take a ten-minute break after two hours of play after which they may continue for an additional hour. (Heo, Oh, Subramanian, Kim, & Kawachi, 2014).

In the Taiwanese city of Tainan, police are encouraged to check for and remove under-eighteens from internet cafés after 10pm (Parkin, 2015, p. 23) effectively enforcing a shutdown law. Taiwanese government have also reportedly passed laws that completely ban children under the age of two from using electronic devices (Locker, 2015). The law classifies electronic devices as being potentially dangerous objects in the same category as alcohol and cigarettes, demonstrating their belief in the link between games, substance abuse, and addiction. Persons under the age of eighteen may not use the devices for unreasonable amounts of time – suggested to be no more than half an hour at a time. Parents not complying with the regulation may incur fines of around $1,500. (Locker, 2015; Seok Hwai, 2015).

We consider the shutdown and cooling off laws to be behaviour policies because they both enforce abstinence for prescribed periods. The laws forcibly remove players from the virtual environment with no consideration for their current in-game commitments. They disrupt standard gameplay practices, potentially forcing players to quit at inappropriate times[2]. Behaviour policies reduce player agency and may cause a player additional stress if they feel they must do the same activities again, or that they have let their team down.

1.2.2 Design policies

The General Administration of Press and Publication (GAPP) in China introduced a policy with the same intention as South Korea's – to limit time spent in game. Much of the available information about this policy is based on news media correspondence. The GAPP recommend that 3 hours is the upper limit of healthy play time, and that after 5 hours, playtime is unhealthy (Yuan & Lee, 2014). However, instead of policing behaviour by strictly controlling the daily play time, the GAPP insist that developers modify their games design, suggesting that developers place a dynamic handicap upon progress (Ernkvist & Ström, 2008).

The Massively Multiplayer Online Role-playing Game (MMORPG) genre is most often associated with problematic game play (Kuss & Griffiths, 2012). MMORPG players must develop their character's abilities to overcome challenges of increasing difficulty. Here, experience points (Exp) are the meaningful unit of character development, and are granted in small increments throughout the game. Players often utilize in-game items such as wearable *gear* that supplements a characters'

[2] For example, when players are far from a save opportunity, or in the middle of a team-based activity wherein other human beings rely on them.

development. These games often have a currency such as *gold* that allows players to buy more gear. For MMORPGs, we refer to the useful rewards of Exp, gear and gold as *loot*. In general, the more time one spends developing a character, the more and better loot that character can access.

The GAPP and Pardo (2010) believe that these loot mechanics promote lengthy sessions, because the quality of loot is related to the total time spent playing. Various sources[3] claim that the policy functioned as follows: players are allowed 3 hours of normal play, but halved the rate at which players gain rewards during hours 4 and 5. The policy completely removed the ability to gain rewards after 5 hours (Ernkvist & Ström, 2008; Yuan & Lee, 2014), potentially by forcing players to log off (Arnason, 2016). For example, if a player ordinarily gains 100% loot per obstacle, then they will gain only 50% loot during hours 4 and 5, and 0% loot after 5 hours, or are forced log off. Players ostensibly recognized that they were becoming less efficient and stopped playing. (Arnason, 2016; Ernkvist & Ström, 2008; Yuan & Lee, 2014). According to one report, the handicap is removed only after five consecutive hours without play (Yuan & Lee, 2014).

1.2.3 Problems with policies

These policies exemplify how action taken to quell the source of a moral panic can be problematic. These policies are fundamentally concerned with reducing screen-time because the authorities believe that screen-time is detrimental and will continue to enforce time-limiting policies despite evidence to the contrary. In October of 2016, The American Academy of Paediatrics acknowledged that time spent in front of digital screens is not strictly harmful and retracted their earlier "nonsensical" suggestion of two hours per day (Ferguson, 2016). Research has shown harm caused by only very excessive viewing (Foster & Watkins, 2010). These policies will likely prevail because it is easier to identify and manage gaming behaviours, such as session length, than it is to identify and treat dangerous social issues such as depression, poor social capital, and decreased life satisfaction. Players suffering from these issues may play games often to offset negative emotions and socialize with other players who accept them, thereby improving their own well-being. Session length restrictions may deny players self-efficacy.

Policies grant adult authorities power over youths, and perpetuate the idea that playing games may be immoral. Authorities may treat players with undue suspicion, and may cause more harm than good (Bowman, 2016). There is anecdotal evidence of players bypassing the behaviour policies that use real-name registration with national IDs to enforce the shutdown, players seek ways to circumvent the restriction by using other people's identities (Abrams, 2010; Bo, 2010) and in doing so they commit identity fraud. Such behaviour at first seems to support the argument that games are problematic. However, Cialdini (2007) argues that *scarcity* is a powerful motivator

---

[3] Many of which reference the same China Daily article (China Daily, 2007)

for behaviour. Cialdini builds upon Brehm and Brehm's (1981) *psychological reactance* theory which has shown that opportunities seem more valuable to people when their availability is limited. In this way, players desire additional time in-game simply because it is denied to them and not because games are addictive. These policies affect all players who engage in those games, or who live in those countries. It is possible that policy damages the medium's ability to generate positive effects by influencing the fundamental experiences that digital games allow.

Games are a crafted medium, meaning that game designers intend to convey specific experiences to their audiences. Yee (2006) found that players also value emergent experiences that result from significant time and emotional investment. Policies such as those identified in Section 1.1 constrict and modify how people play and may meaningfully influence the player's experience of the game. The strict policies employed by South Korea and Taiwan directly manipulate session length, shutting down a session before the negative effects presumably manifest. The design policies dynamically manipulate the mechanics of a game, fatiguing character growth and stunting player progression. This strategy attempts to change the subjective experience enough to make a player choose to end their session early. It is not clear how this policy affects positive subjective user experience, or session length. Such practices potentially change the subjective user-experiences. If these policies do change user-experience significantly, they may degrade the fundamental qualities of the experience and negate positive effects such as their ability to offset negative moods, and foster meaningful social relationships.

1.3 Research questions

The nationally institutionalized policies identified in Section 1.1 affect the entire games industry as well as the governing bodies that implement them. They affect players every time they play because the systems are permanent. They affect the commercial success of games that rely on positive experiences, or on in-game revenue from purchases or adverts. The design policies affect developers when the games require special systems or mechanics that were not a part of the initial game design. The governing bodies that implement and enforce the policies require resources to ensure they run as intended. Given the relatively low incidence of negative outcomes in games, it is unclear if society needs policies such as these, if they are cost effective, and how they affect players in terms of their subjective experience and motivation for future play. If players regularly play for the full duration that policy allows, then their experience may deviate from those intended by the developer, and those that emerge because of time investment such as relationships. This dissertation seeks to investigate these uncertainties by answering the following questions:

1. How do design policies change a player's session length?

The literature alludes to session length and screen time as a predictor, and sometimes cause, of negative effects. The amount of time spent playing games appears to be the primary factor that the policies seek to influence. We know the behaviour policy must affect session length in some cases because it *is* a time limit. The design policy grants players unlimited session length but limits the period in which players can effectively develop their character. For these players time is again a limited resource. In all cases, session length is a limited resource; players are unable to access the intended experience beyond a restricted period. If the design policies do not decrease session length, then their only outcome is to limit experience.

2. How do the behaviour and design policies change a player's subjective user-experience?

The behaviour policies provide players with a fixed period of play across all games. They do not take the gameplay cadence of individual games or genres into account. Players who wish to play for the full period may find themselves in the middle of an intense passage of play when they run out of time, forcing them to end their session. For single player games, they may lose progression if they were unable to save their state. For multiplayer games, they are forced to abandon their peers and risk rebuke or punishment from game marshals or punishment systems. In these cases, policy denies players experiences such as satisfaction, autonomy and a sense of accomplishment. They further put players at risk of unintended negative experiences, such as frustration or anger. The design policy appears to recognize these risks and allows players to continue. However, the implementations specifically modify game mechanics and rewards, changing the nature of the game (at least in part) and possibly affecting the experience.

3. How do the behaviour and design policies change a player's intention to return to a game?

Player experience influences a game’s success, suggesting that players do not intend to play in the future if they have a bad experience, or have positive experiences removed. The policies of concern are an attempt to change how the player manages their session length, but it is possible that the policies affect when players plan to start playing again. For example, the experience of a policy may increase a player’s desire to remain in the real world rather than the virtual, which reduces total time spent in game. However, if policy increases desire to play, then players may seek illicit means to access the game, such as identity fraud (see Section 1.2.3).

4. How does the design policy change a player's reason for choosing to end their session?

The design policies restrict character progression which is an important mechanic in RPGs, and at least part of the reason players choose these games. It is unclear if players do stop after their character progression is fatigued, and whether they stop playing because of it. Players may continue because they invest in the story rather than character progression, or continue because they are having more fun socializing than playing with the mechanics. The behaviour policies remove player agency in how they choose to end their session should they use all their allotted time, and is therefore not considered in this question.

1.4 Outline

The body of this dissertation seeks to answer the questions formulated in this chapter. In our Background chapter which follows this section, we discuss theories and models of subjective user-experience in digital games and operationally define the construct this research. In our Methods chapter, we discuss the user experiment we designed to help answer our research questions. We detail our choice of method, before describing the experimental system we built and our choice in metrics and finally describe our data collection practices. Our Results chapter describes our sample, the data they generated, and the relevant statistical analyses. Our Discussion chapter draws inferences about the user-experience from our analyses. Our Conclusion chapter situates our findings within the broader contexts of games research, policy-making, and games design. Following this are all appendices that may prove useful or informative to the reader.

2 BACKGROUND

2.1 Introduction

In the preceding chapter, we discuss how digital games affect players in a variety of ways, both positively and negatively. In this chapter, we elaborate on the effects of digital games. We then comment on the complexity of studying digital games as a medium with a brief history of games studies and explain why we focus on the experiential phenomena associated with games for this research. We end with a review the prominent theories and models associated with subjective user experience in digital games (hereafter referred to as *UXG*), and discuss their relevance to our research questions.

2.2 The effects of digital games

2.2.1 A healthy hobby

Digital games have the potential to improve well-being in a variety of ways. Some are effective teaching tools that let players hone domain-specific skills such as typing or arithmetic, whereas others have positive effects that improve human function and bolster well-being (Bowman, 2016; Connolly, Boyle, MacArthur, Hainey, & Boyle, 2012; Granic et al., 2014; Vella & Johnson, 2012). For example, the same types of violent games often associated with increased aggression are positively correlated with improved visual-spatial abilities (Ferguson, 2007). Action games teach players to allocate attention more effectively than non-gamers, a useful skill in every-day life (C. S. Green & Bavelier, 2012). In a study of socio-emotional states of older adults (average age 77), the 60% who self-identified as gamers reported greater levels of subjective well-being and social functioning, and further displayed less negative-affect[4] and fewer signs of depression (Allaire et al., 2013). This is because the regular experience of positive emotions improves subjective well-being by affirming their ability to experience positive emotions; they expect positive experiences in the future and thus perceive their lives to be good (Fredrickson, 2013).

Players use games to manage their mood, especially to offset so-called "bad" moods (Olson, 2010; Olson, Kutner, & Warner, 2008; Ryan et al., 2006). Granic and colleagues (2014) suggest that digital games are among the most effective of mood management tools because they generate many positive experiences, including a flow state that has long been associated with well-being (Csikszentmihalyi, 1990). Regular flow experiences motivates people to overcome challenges at their perceived limit, thereby affirming their abilities to overcome challenges in general (Moneta, 2004). This promotes self-confidence and self-esteem, and further encourages the pursuit of "lifelong organismic growth." (Moneta, 2004, p. 116). In this way, digital games offer an efficient means to improve well-being.

Youths use digital games to meet a variety of needs ordinarily provided by other social situations and settings, but with far less risk attached (Kowert, 2016). Some allow players to choose their avatar's age, gender, appearances and body type. This affords players the ability to experiment with self-identity and identity expression in a low-risk environment (Granic et al., 2014). Digital games can afford players the opportunity to explore socially unacceptable practices that include both generally delinquent acts such as vandalism or violence (Coulson & Ferguson, 2016; Olson, 2016), and sometimes socially unacceptable behaviours such as non-normative sexuality (Hussain & Griffiths, 2008). Massively multiplayer games allow players to interact in ways that require cooperation and competition. They provide the same opportunities as organised sports but are affordable and players are far less restricted by physical resources, schedules, and team management[5]. This provides players with many opportunities to test the limits of important and sometimes tricky life skills such as setting goals, frustration management, and self-promotion (Adachi & Willoughby, 2013; Olson, 2016).

2.2.2 An unhealthy vice

There have been concerns regarding the behaviours and health of digital game players since the medium's inception. Soper and Miller (1983) offer some of the earliest reports of problematic behaviour in adolescent gamers describing the activity as an addiction. Kuczmierczyk (1987) offers reported successful rehabilitation of a patient that spent too much time, and money, on coin operated arcade games. Reasoning that the problem was like compulsive gambling behaviours, the authors used

[4] Observable physiological manifestation of negative emotions. See Section 2.4.3

[5] Organized sports activities can be subject to limited track/field/court availability, where leagues and matches are scheduled months in advance. MMO games on the other hand are less bound by physical limitations, and organizing a match and finding teammates can be as simple as clicking a button.

contemporary addiction treatment methods. Discussions about video game addiction persist today, alongside additional concerns about aggression, social development, and health. Digital games and excessive play are implicated in many publicized texts including a judge's sentencing of a murder suspect that explicitly linked the crime to digital games (Burge, 2009). News outlets commonly report on a range of tragedies such as an adult assaulting a child after their in-game interaction (Fleming, 2011), players dying at their computers during a marathon session (Prynne, 2014), and the motives behind mass murderers (Pidd, 2012; THR Staff, 2013). Such news media paint a bleak picture of games and their risks.

These effects certainly deserve academic attention. As the body of literature grows, we begin to understand that problematic behaviour is relatively rare (Ferguson, 2015; Gunter & Daly, 2012; Kowert & Quandt, 2016; Przybylski, 2014). Recent surveys of players that meet the diagnostic criteria for the proposed *Internet Gaming Disorder* (American Psychiatric Association, 2013) report that less than one percent experience any significant negative effects at all (Przybylski, Weinstein, & Murayama, 2017). This suggests that digital games are not as problematic as once thought. Where violence and aggression concerned, the literature is divided. Games researchers found evidence suggesting that digital games increase aggression, and evidence suggesting no effect (Kowert & Quandt, 2016). Meta-analyses suggest that the relative impact of violent digital games to society is minimal. Researchers have identified methodological inconsistencies across aggression literature, most notably issues of measurement and conclusion[6] (Coulson & Ferguson, 2016).

2.2.3 Effects in perspective

Games appear to be generally affecting. The research shows that players receive myriad benefits by playing digital games, but also are at risk of suffering some disadvantages. They help players improve a variety of domain specific skills, promote positive emotions while offsetting negative emotions, and provide safe spaces that allow experimentation and personal growth. These effects are desirable, and while they are not unique to the medium, games are popular and compelling which may increase their efficacy over other media and practices. This is not to suggest that games are a cornucopia of benefits, nor a panacea for common problems. These effects are not without caveats. If players use digital games to avoid confronting negative emotions entirely then play can become maladaptive as players do not learn to take action to manage their problems (Granic et al., 2014). For example, competitive gaming may foster maladaptive self-aggrandizing, increase toxic and harmful behaviours, and reduce empathy for others if player reactions are left unmediated (Olson, 2016).

---

[6] For example, one study measured aggression based on how much hot-sauce one player would serve to another (Lieberman et al., 1999).

The literature certainly suggests that games may negatively affect certain individuals. But the quality of well-known research is sometimes problematic. For example, Soper and Miller (1983) chose to use their subjects' own colloquial language to describe their observations: the terms *addiction* and *junkie* which are metaphors laden with moral connotations and tend to inspire value judgments. Lieberman and colleagues' (1999) hot sauce allocation experiments make for fascinating reading, but their metaphor for aggression does not compare to gun-violence and murder; short-term increases in aggressive thoughts do not translate into a long-term increases in aggressive behaviours (Kowert & Quandt, 2016).

Contemporary research that attempts to identify and define gaming related disorders is rigorous and sophisticated, though it continues to makes use of the addiction and substance use metaphors despite a wealth of criticism regarding the conceptual differences between terms like addiction and the disorders in question (Aarseth et al., 2016; Kuss, Griffiths, & Pontes, 2017). Such metaphors are too presumptuous to allow critical use or generalization. For example, a troubled individual who happens to play games should not be diagnosed with a gaming disorder simply because they play many games if a different and better understood disorder is responsible for their troubles.

Information regarding prevalence is relatively new (published in the late 2000s and early 2010s) but it shows that the risk that players will experience problematic gaming outcomes is low. Nonetheless, the notion that games are harmful and must be strictly moderated prevails and must be considered until we are certain that they are not (Lorentz, Ferguson, & Schott, 2015). Rather than blame the medium, research should strive to identify the factors that precipitate each effect (Kowert & Quandt, 2016), and practice methods that promote the positive aspects while minimizing the negative aspects.

2.3 Digital Games: A complex entertainment medium

Player focussed digital games research traditionally asks if games themselves cause a particular effect (Kowert & Quandt, 2016). Such questions are fraught with problems because they misconceive the functional and formal differences between individual games. The term *digital games* encapsulate a plethora of styles, genres and activities, each taxing the mind and body in such different ways that we cannot predict the effects without knowing the details of a game, who is playing it, and how they are playing. By analogy: *Farmville* (Zynga [Video Game], 2009) is to *StarCraft* (Blizzard Entertainment, 1998) as walking is to formula racing. In this section we briefly look at definitions for digital games, and the history of games research.

2.3.1 A brief overview of the definitons and history of games research

"Playing a game is the voluntary attempt to overcome unnecessary obstacles." - Bernard Suits (1978, p. 41)

The above definition may be succinct, but games and play are two complex and interrelated concepts that describe the traditions and behaviours of individuals and groups observed in human and animal life. Games and play have myriad definitions which describe these phenomena's importance to individuals and the societies in which they live. Salen and Zimmerman (2003) review eight considered definitions including those by well-known scholars of play and games Caillois (1961), Huizinga (1949) and Suits (1978). By comparing these, they define a game as "a system in which players engage in an artificial conflict, defined by rules, that results in a quantifiable outcome." (Salen & Zimmerman, 2004, p. 65). This definition encapsulates both the game *object* and the *play* activity within a *system* thereby emphasising their interdependence.

The definitions presented above describe all types of games including the digital, or video, formats of interest to this research which have their own additional complexities. For some time, scholars debated over how to approach digital game studies. The *narratologist* argued that digital games should be studied as new forms of text, similar to literature and drama, that facilitate story-telling (Frasca, 1999) whereas the *ludologist* viewed digital games as a system of rules or mechanics facilitating the action of play (Frasca, 1999; Salen & Zimmerman, 2003). The debate produced no winner; Murray (2005) suggests that neither approach is ideal because both qualities are increasingly present in individual titles. The complexity of the medium is well described by Ivory: "This kludge of technological and social bloodlines and audiences under the loosely defined blanket term *video games*" (2016, p. 2).

To elaborate, Ivory (2016) traces two ancestries common to all modern video games and effectively justifies both the narratology and ludology perspectives, although it does not appear to be their intention. They describe *shared narratives* heritage inspired by the fantasy archetypes of Tolkien's Middle-earth[7], and the thematically related table-top games such as Dungeons and Dragons[8]. Digital games of this ancestry act as a medium primarily for the creation and telling of stories; something. Ivory (2015) also describe *action simulation* games that abstract to computers real-world examples of competition and combat, such as sports and war. Digital games of this heritage provide safe environments for players to take part in the simulated activities governed by rules and mechanics. Many modern digital games draw from

---

[7] Tolkien JRR. *The Hobbit: Or, There and Back Again*. London: George Allen and Unwin;1937. Tolkien JRR. *The Lord of the Rings.* London: George Allen and Unwin; Part one 1954; Part two 1954; Part three 1955

[8] Gygax, G. and Arneson, D. Dungeons and Dragons. Lake Geneva, WI: Tactical Studies Rules; 1974

both ancestries, the most successful of which foster a symbiotic relationship between action elements and narrative elements to generate ludo-narrative experiences for the player, where the content is inextricable from the act of playing. A holistic view such as this helps to differentiate games from other forms entertainment media, and suggests that games research requires a careful approach.

2.3.2 Approaching games research

Kowert and Quandt (2016) call for a nuanced approach to games research that considers the variations in *genre* and *game*, and *players* and the social *contexts* in which they play. They note that games research often relies on simplified models of these four aspects to identify effects, which is then generalized to all games and players. For example, addiction research focuses on the Massively Multiplayer Online Role-Playing Game (MMORPG) genre of games (Griffiths, 2016; Kuss & Griffiths, 2012; Metcalf & Pammer, 2014); the observed effects should necessarily be attributed to the First-Person Shooters (FPS) genre, as the player's experience differs due genre conventions including mechanics and progression. Individual games within genre itself provides limited reliability. Aggression research focuses on FPS games because they tend to depict gun violence[9], and because they offer more scope for experimental control. Thus, we know to be careful when selecting an appropriate game for our experiment (see Section 3.3.1) and discussing the implications of our results (see Sections 5 and 6).

Games discourse risks generalizing that all people play for the same reasons, in the same way, and are affected similarly. Such a view does not account for why some players display a dependence on a specific game while most others do not (Kowert & Quandt, 2016). It is clear that individual personal differences dictate how players spend their time playing games, and how they interpret and respond to events while playing games (Allahverdipour, Bazargan, Farhadinasab, & Moeini, 2010), ultimately affecting their experience. Researchers should be sensitive to these differences to avoid generalisations. Games researchers sometimes prefer formal classifications to guide interpretation of individual differences. Some use trait-based theories of personality, such as the predominant descriptive Five-Factor Model (John, Naumann, & Soto, 2008), to predict a variety of motivations to behaviours such as play styles (Bean, Ferro, Vissoci, Rivero, & Groth-Marnat, 2016; Bean & Groth-Marnat, 2016), enjoyment (Fang & Zhao, 2010), game preferences (Johnson & Gardner, 2010), specifically towards violent games and play (Chory & Goodboy, 2011). Advances in neuropsychology allows researchers to use models such as Reinforcement Sensitivity Theory (Gray, 1970; Gray & McNaughton, 2003) to investigate how individuals respond to rewards presents in games.

[9] Although some depict no violence whatsoever; thus, the mechanisms purportedly leading to aggression may be absent.

Whereas the previous models are general to all behaviours, games researchers also seek to classify player differences according to more domain specific, or genre and game specific, criteria. Bartle's (1996) player typology may be the most well cited example. It describes four motivating playstyles differentiated by what it is that players spend most of their time doing in a Multi User Dungeon game. Researchers should be cautious extending this to other games and genres. Lazzaro (2004) argues that games generate emotions, rather than activities, and describes four motivating playstyles that lead to specific experiences. For a thorough discussion and meta-synthesis of player typologies, see (Hamari & Tuunanen, 2014). Banks and Bowman (2016) argue that the connection between player and avatar produces cognitive, affective, and behavioural phenomena which are crucial to player experience. They integrate two models of this connection to develop a typology to describe *player-avatar relationships.* Some players view avatars as tools, whereas others engage with their avatars in ways that resemble human social relationships (Banks, 2013). Researchers should consider using such paradigms if their research questions require this type of deliniation; if they aim to identify patterns in large samples of players via survey or investigate classification as means to make predictions about players themselves. Although our research aims to compare policies, not players, we must be aware of individual differences. See Section 3.2 for discussion on methods, and Section 3.2.2 in particular for working with extraneous player variables.

Personal factors alone do not account for the effect variations People play digital games in many social contexts. Players can play single-player or multi-player games, online or offline, and while physically isolated or while physically co-located with other humans. Lenhart and colleagues (Engeser & Schiepe-Tiska, 2012) report that only 24% of all teenagers surveyed exclusively play single-player games, and that 71% of teenagers in the United States play in a variety of other social contexts. Interestingly, research on *World of Warcraft* (Blizzard Entertainment, 2004) revealed that most players spend less than a third of their play time with other players (Przybylski, 2014), calling into question the appeal of mass socialization in these games. Different social contexts place different demands on a player's cognitive resources, generating different experiences. For example, the mere presence of an audience can improve player performance (Engeser & Schiepe-Tiska, 2012). In this and many other ways, the *context* in which the same digital game is played directly influences the experiences that a player has, and the effects of the game on the player.

The policies identified in Section 1.2 likely interact with the dynamics of genre, game, player and context. Specifically, the behaviour policies (see Section 1.2.1) restrict the context of play; players are limited by when, and for how long they can play. The design policy (see Section 1.2.2) utilizes characteristics of the genre, such as loot gain, and is necessarily bespoke, at least in implementation. It stands to reason that policy will influence player experiences in some way. To identify changes in experience, we must understand how experience manifests in games.

2.4 Models and measures of experience in games

2.4.1 Introduction

Salen and Zimmerman's (2003) definition for digital games includes a *quantifiable outcome,* which explicitly refers to the completion of the story arc for narrative games, or a win-lose condition or score for action games. Games are a crafted medium, meaning that designers intend for players to have specific experiences and so the definition implies experiential outcome as well. Furthermore, the experience tends to be desirable, if not inherently joyous. Games are "just for fun" because the pretend *conflicts* that games create is *artificial*; defeat is not permanent. Fun is a well-known word for a complex idea that carries connotations of pleasure, amusement, and enjoyment. Huizinga (1949) considered fun to be a vital human experience because the regular experience of fun has lasting positive effects on subjective well-being. Playing digital games is fundamentally a positive experience but the psychological experiences are as diverse and fantastic as the games in which they occur (Cairns, Cox, & Nordin, 2014).

Time limiting policies *potentially* negate these rewarding experiences by fatiguing, and changing, the mechanics[10] of a game, or how players can interact with a game. Our research seeks to understand how such policies affect the user experience in digital games. To qualify the unique experience of playing digital games, players, reviewers and developers adopted several domain-specific terms, such as a colloquial definition for *immersion*. Furthermore, researchers wished to quantify player experience, and looked towards established models of human experience such as *enjoyment, engagement, flow, presence,* and *affect,* or operationalized the colloquial notion of immersion specifically for games (Nordin, Denisova, & Cairns, 2014; Norman, 2013; Sweetser & Wyeth, 2005). In the following section, we express a current understanding of user-experience in the digital games domain (hereafter referred to as *UXG*).

More than fun

That digital games are meaningless fun is an increasingly acknowledged misconception (Oliver et al., 2016). Players experience the variety of emotions associated with all forms of goal-oriented activity, including traditionally undesirable emotions such as frustration and disappointment (Lazzaro, 2004; K. Poels, de Kort, & Ijsselsteijn, 2007b). Furthermore, many games intend to induce uncomfortable feelings. Hotline Miami (Dennaton Games, 2012) is framed within an exhilarating aesthetic, and has players murder masses of Russian mafia through (admittedly) fun game mechanics. However, before each level completes, the game interrupts the

---

[10] The rules of the game, and the mechanism through which a player acts within the game world.

mechanic and fast paced music and intentionally generates cognitive dissonance[11]. Players recognise that the game's mechanic is fun, but the game confronts them with the hyper-violence that results in enacting its mechanic. In doing so, the game provides meaningful experiences in addition to fun. Amnesia: The Dark Descent (Frictional Games, 2010), much like horror films, purposefully generates feelings of distress (including horror, fear, and sometimes nausea) by threatening players with unknown virtual peril. Life is Strange (Dontnod Entertainment, 2015) allows players to affect the outcome of a poignant narrative – players commonly report powerful emotional reactions. All three examples above have achieved commercial success without solely relying on fun. Fun remains a valid but incomplete description of digital games' various experiences. Knowing this, researchers seek to describe said experiences with more appropriate nomenclature.

Enjoyment and other experiences

Digital games can also be considered to be an entertainment medium (Cairns et al., 2014). In those terms, research into the experience of digital games borrows terminology and direction from established fields of entertainment and media psychology. Entertainment media discussion has focused on *enjoyment* as the desired response, which encompasses *fun, amusing* and *thrilling* experiences (Oliver & Bartsch, 2010) and has been defined in a variety of ways (Tamborini et al., 2010). Where digital games are concerned, researchers go as far as to call enjoyment "the single most important goal for computer games." (Sweetser & Wyeth, 2005, p. 1544). However, *enjoyment* carries connotations of pleasure seeking that preclude discussions of unpleasant but rewarding experiences (Tamborini et al., 2011). For example, an audience may not consider a documentary depicting genocide or slavery enjoyable, but they enjoy *having watched* the film because they can derive and appreciate meaning from the content or form regardless of how bleak the content may be (Oliver & Bartsch, 2010). More generally, the desired outcome for entertainment media and digital games is a *holistically* affecting experience, not necessarily fun, whereby the experience of emotions adds meaning to the activity.

Enjoyment, as a construct, offers a more nuanced description of experience, and although it acts as a catch-all term for the positive experiences in entertainment media, UXG requires nuance, and uses well established theories of subjective experience for its definitions. The are a large number of models and measures commonly used in UXG research. In their paper *Too Many Questionnaires: Measuring Player Experience Whilst Playing Digital Games,* Nordin and colleagues (2014) identify

---

[11] The state of having inconsistent thoughts, feelings and attitudes, especially relating to recent behaviours – verbatim from internet. Alternatively, a human holds two conflicting ideas, thoughts or beliefs simultaneously. In this case, they hold that the gameplay is fun, but are also confronted with the reality of violence.

seven notable models of UXG based on the availability[12] of a relevant questionnaire that allow direct measurements of player experiences [13]. These are *flow* (Csikszentmihalyi, 1988), *presence* (Witmer & Singer, 1998), *immersion* (Brown & Cairns, 2004; Jennett et al., 2008), *GameFlow* (Sweetser & Wyeth, 2005), *game engagement* (Brockmyer et al., 2009), *player experience of needs satisfaction* (Ryan et al., 2006), and *social presence in gaming* (de Kort, Poels, & Ijsselsteijn, 2007), which is one module of the G*ame Experience Questionnaire* (Ijsselsteijn, de Kort, Jurgelionis, & Bellotti, 2007; K. Poels, de Kort, & Ijsselsteijn, 2007a; K. Poels et al., 2007b). Ironically, *Too Many Questionnaires* fails to review all the available UXG questionnaires, and reintroduces the *Flow Questionnaire* to the discourse, which has known weaknesses (Moneta, 2012), and there is good reason to doubt that the Presence Questionnaire has construct validity (Slater, 1999). Wiemeyer and colleagues (2016) provide a review of some additional models, including two measures of spatial presence (Hartmann et al., 2016; Vorderer et al., 2004), a model describing the *Core Elements of the Gaming Experience* (Calvillo-Gámez, Cairns, & Cox, 2010), and two models discussing enjoyment in serious games (Fu, Su, & Yu, 2009; Wouters, van Oostendorp, Boonekamp, & Spek, 2011). In addition, there exist other significant applications of flow, including Jones's (1998) adaptation of the flow components for game-based learning and Nacke's (2012) reduction of four flow models into four components. In the following sections, we review the most prominent models of player experience.

2.4.2 Flow

Flow is a positive state of mind that describes the universal subjective experience associated with performing activities (Csikszentmihalyi, 1990). Sometimes called optimal experience (Cowley, Charles, Black, & Hickey, 2008; Csikszentmihalyi & LeFevre, 1989) flow is well suited to describe the experience of playing games. It posesses strong and positive correlations with motivation to perform an activity, feelings of competency, as well as lasting happiness and well-being (Engeser & Schiepe-Tiska, 2012). Flow is an intuitive concept. An athlete in competition state might say they were "in-the-zone". A musician practicing a difficult piece may be oblivious to their spouse asking them a question. Players in an hour long Dota 2 (Valve Corporation, 2013) match may feel that only twenty minutes have passed. Such are colloquial descriptions of a flow experience (Csikszentmihalyi, 1988; Engeser & Schiepe-Tiska, 2012). Formally, flow describes a person's mental state when they engage so wholly with a task that they lose awareness of events unrelated to the activity, such as sense of self and outside stimuli.

[12] Strictly speaking, only six of these seven models possess an accessible questionnaire. *GameFlow* is a set of heuristics for evaluating game design, not a questionnaire evaluating individual player experience.

[13] The authors use *engagement* interchangeably with *experience*

Table 2.1 Nine Components of Flow and how they manifest in games. Adapted from Engeser and Schiepe-Tiska (2012), and corresponding Game-play Elements, adapted from (Cowley et al., 2008).

A task that requires skill and mental energy described by flow components	A game made up of game-play elements
Focused attention on the task at hand.	Sensory immersion and distraction-free environment dedicated to gaming.
Strong sense of control over the task's demands.	Familiar interfaces and game-play mechanics; Agency and confidence for success.
Action merges with awareness. (Thought is action; there is no conscious mediation).	Sense of non-mediation by computer in the virtual environment, and mastery of game mechanics.
Intrinsic motivation - the doing motivates action, not the completion of the task.	High motivation to play, no imperative to do otherwise; empathetic to content.
A loss of reflective self-consciousness.	Embodiment in game avatar.
A warped sense of time passing – chronoslip (Parkin, 2015)	Focusing on another, temporally independent environment.
The task has clear, unambiguous rules and goals.	Missions, plot lines, levels; any explicit outcome during a successful play session.
Immediate feedback on actions.	Well-timed, suitable rewards and penalties.
Tasks perceived to be challenging but feasible.	Game design elements such as difficulty levels and scaffolding

In Table 2.1 we describe flow by nine components (Engeser & Schiepe-Tiska, 2012) and their applications to games (Cowley, Charles, Black, & Hickey, 2008).

Models of flow are important to player experience for two main reasons. Firstly, all flow components describe positive aspects of player experiences (such as fun, enjoyment, and entertainment) and the components of flow map easily to the elements of digital games (see Table 2.1). For this reason, flow is one of the most common experiential models used in digital games research. Secondly, in recognition of this affinity, developers consciously design their games to facilitate flow. Flow is a useful tool for understanding individual player experiences. In the following paragraphs, we briefly review some theory of flow, as well as the pertinent applications to games.

Research suggests that no single component should be enough to describe a flow state, but that not all components need to be present for flow to manifest. It follows that flow is the interplay of its components. Furthermore, some components may act as pre-conditions for a flow state, rather than result from a flow state. An environment dedicated to gaming can encourage flow by minimizing distraction, whereas a player already in flow does not notice distractions. (Engeser & Schiepe-Tiska, 2012).

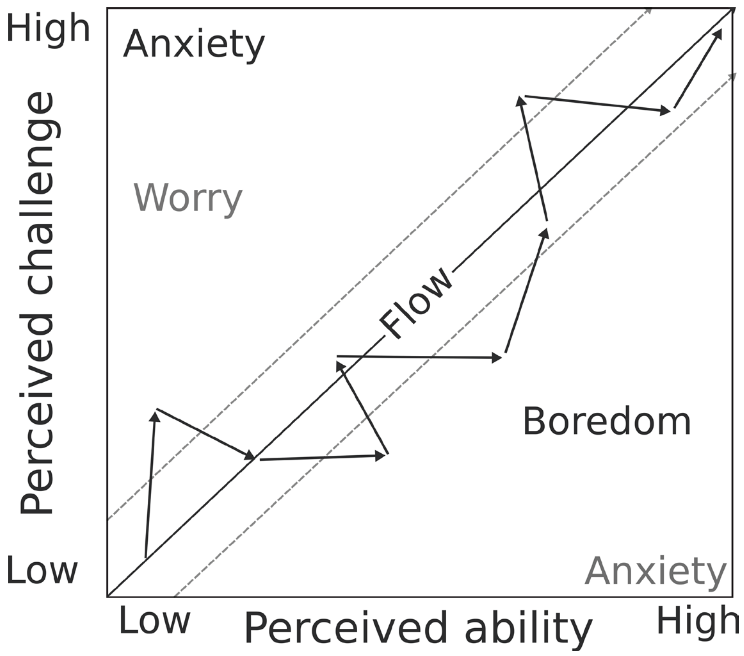

Figure 2.1 The so-called "First" model of flow state redrawn from Csikszentmihalyi.(as cited in Moneta, 2012, p. 26) Flow occurs when an actor perceives that their ability matches the challenges of an activity. The arrows represent how an activity may change, and the actor's experience may fluctuate as their perceptions of it change.

Despite the multi-dimensional nature of flow, the literature often focuses on a subset of components when applying flow to specific activities. This reduces the complexity of discussion and focuses on the components that have the largest effect on flow for the given activity. With games research, the warped sense of time passing that Parkin (2013) names chronoslip receives some attention because it is common to lose track of time while playing, especially when having fun. Chronoslip is also implicated in problematic gaming behaviour; by warping the sense of time passing, the experience steals time from players if they accidentally play longer than intended (see *The dark side of flow* p. 22). Games design and experience literature focuses on the perceived balance between challenge/demand and skill/ability (Løvoll & Vittersø, 2014), and sometimes equates the component to flow (Chen, 2006; Csikszentmihalyi & LeFevre, 1989; Eisenberger et al., 2005; Hektner & Csikszentmihalyi, 1996; Nacke & Lindley, 2008; Qin et al., 2007). Such treatment is likely due to the relative ease with which researchers can manipulate challenge. Any given activity presents a specific amount of challenge and any given actor has a unique level of skill (or set of abilities) with which they tackle that challenge.

Many behavioural researchers attempt to facilitate this balance to better engineer flow states. Early graphic models portray flow as a sector within an abstracted emotional spectrum, where flow exists somewhere between boredom and anxiety (see Figure 2.1 p. 20). An actor's perceived balance between these two factors is far more important than the objective levels (Csikszentmihalyi & Massimini, 1985). As a consequence, the position of any given actor's flow sector is entirely subjective. This balance component is especially important in digital games research and development where challenge is a crafted and malleable game element and key to the fundamental experience: to control the balance between challenge and skill is to ensure flow, and guarantee a positive game experience with lasting positive consequences.

Flow in games

Chen's (2007) application of flow to games may be the most popular, if not the most successful. Chen acknowledged that a sense of control in a flow state is generally considered to be granted by virtue of being in that state, often because actors can use their skill to overcome a challenge. Their insightful contribution suggests that games should grant players continuous autonomy over the degree of challenge presented by the game at each obstacle. This approach celebrates the important *perceived* aspect of the challenge-skill balance, and effectively grants players another component of flow (in this case, control of the game's difficulty) so that they themselves may find balance. Their method uses an analogy[14] to conceptualize how control is granted to participants in games: Consider a group of people practicing high-jump. A horizontal bar allows only those who can clear the current height to succeed and continue. All those who cannot clear the bar are eliminated. Slanty Bar method suggests that the bar be set diagonally, so that all jumpers may choose their challenge height for each jump. Actors that seek a challenge can push their limits, whereas actors that seek competency can choose a comfortable height. In this way all jumpers may participate all the time. As applied to games, Chen proposes that before any obstacle is attempted, players should be allowed to choose an approach that allows them to match their skills to it (see Figure 2.2). Whereas being a in a state of flow grants players a *sense* of control, Chen suggests that designers facilitate control to make flow more accessible. (Chen, 2006, 2007). As a result, many games companies now try to build flow into their games. Thus, games intentionally offer players far more agency since Chen's work, and flow has become intergral to the study of UXG. That said, we must still bear in mind that the challenge-ability balance may be important to the flow experience, but one must remember the other components, and not reduce the construct to this simple ratio (Landhäußer & Keller, 2012).

[14] Musston's Slanty Bar analogy is an intuitive idea that is often referenced but never cited. Chen references a discussion of the analogy in Dekoven's work Of Fun and Flow (DeKoven, n.d.))

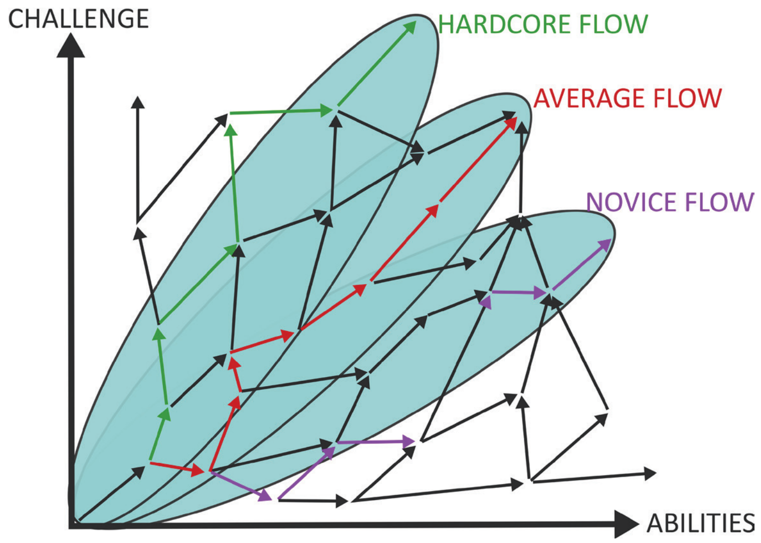

Figure 2.2 Chen's adaptation of the slanty bar to difficulty in games.Arrows represent a path chosen by the player at any challenging obstacle. Black arrows represent all possible paths. Red represents the path an average player might take to stay within their subjective flow zone. Purple and Green represent paths taken by novice and hardcore gamers respectively, to stay in their respective flow zones. Adapted from (Chen, 2007).

Cowley and colleagues (2008) adopt a theoretical approach, restructuring flow components for digital games (see Table 2.1 p. 19). Specifically, they model gaming as an information system made up of user (player), system (game and genre), and experience (flow). They argue that new information, such as novel game elements and challenges confuse players, and creates in them a state of *psychic entropy* (Cowley et al., 2008, p. 24). Players may enter a flow state while processing the new information and try to restore cognitive order. Once players complete the process, they successfully apply their strategy and progress in the game, where new information presents itself, thus creating additional psychic entropy. This model essentially describes how perceived challenges and abilities manifest between game and player during the activity. It further presents the notion that flow intensity fluctuates throughout the game depending on the amount of psychic entropy the game creates. However, this approach reduces players to simple information processors that display their states and nothing more. This trivializes the extent to which individual human variations affect the experiences.

The dark side of flow

Flow is considered to be a highly rewarding state of mind and an experience that improves subjective well-being. However, there exist notions that flow can have (rare) negative or detrimental consequences (Schüler, 2012). People may actively seek a state in which they "lose the ultimate control: the freedom to determine the contents of consciousness" (Csikszentmihalyi, 1990, p. 62) in order to gain the reward. This may place one's well-being at risk, as they potentially cannot attend to themselves or their immediate environment, a situation common in problematic video game playing (Schüler, 2012). There is little doubt that digital games facilitate flow experiences, and similarly little doubt that digital games facilitate bahaviours that are detrimental to quality of life. These two phenomena may not be causally related, but some authors suggest that they may be associated.

Schüler (2012) draws parallels between flow characteristics and the International Classification of Diseases' definition for dependence syndrome, while others look at the role that flow plays in problematic gaming specifically. For example, that the experience of flow is a strong motivation to continue returning to a certain game experience (Qin et al., 2007) and that problematic behaviour is formed when oft-repeated activities result in a state of flow (Chou & Ting, 2004). Other explorations of problematic video game playing address issues that might be explained through the lens of flow. Wood (2008) identifies poor time management as a reason for excessive game play. Players cannot apply time management skills while in a flow state because of the warped sense of time passing therein. Thus, players may lack the agency to decide to stop playing. This is not apparent in activities that have a behaviour-control strategy external to their mind, such as sports with strict limitations on game time. In these instances, time is mediated - the rules of these games define the beginning and end. In digital games such as MMORPGs, such rules do not exist. The virtual world is persistent. Standard restrictions on time such as day and night, as well as related physiological restrictions, such as circadian rhythms, often hold little consequence in the game world; a character's fatigue, for example, is treated with magic. Thus, these limits on the player's avatar can be ignored or dispelled. Games have been shown to increase adherence to cancer treatments (Kato, Cole, Bradlyn, & Pollock, 2008) and virtual environments can distract patients during painful burn rehabilitation procedures (Hoffman et al., 2011). Players are likely to be distracted from discomforts such as fatigue and hunger in the same manner (Griffiths, 2005). The policies (see Section 1.2) acknowledge this idea by introducing control systems that do not rely on the player's will power.

There have been no studies that have found flow to be the cause of problematic gaming at the time of writing. Indeed, flow is widely regarded as a positive and universal experience. In contrast, problematic behaviours are so rare that flow's utility is almost without question. It is nonetheless important to consider the role that such experiences might play in problematic gaming, especially when designers create their games with the intention of inducing flow states. If keeping a gamer in a flow state for too long is bad, then the flow state should be intentionally interrupted. The behaviour and design policies may already do this by intentionally breaking the player's flow experience.

2.4.3 Emotions and affect

"What makes failing 80% of the time fun?" (Lazzaro, 2004, p. 1)

The act of playing digital games provokes many emotions in its players. Over and above the emotions that storytelling can evoke, Lazzaro (2004) identified 30 different emotions that are provoked by game-play alone. Emotions influence our subjective perceptions and reactions to events in the world (Riva et al., 2015), they play a role

Table 2.2 The emotions representing the Positive and Negative Affect Schedule (D. Watson, Clark, & Tellegen, 1988).

Positive-affect	Negative-affect
Active	Afraid
Alert	Scared
Attentive	Nervous
Determined	Jittery
Enthusiastic	Irritable
Excited	Hostile
Inspired	Guilty
Interested	Ashamed
Proud	Upset
Strong	Distressed

in the incidental and prevailing effects generated by digital games, and certainly guide the motivations for why gamers seek the activity.

Affect is a psychological term that describes both the physiological presentation of emotions, and the subjective psychological experience of emotions. For example, when a person is enjoying something, they spontaneously smile and have a pleasant feeling. A succinct definition of the construct is difficult due to the complexity of the phenomenon it describes (Kelly, 2014). It can be thought of as the resulting emotional output from a given life event input. In context of digital games, affect is the subjective emotional experience (output) that is the result of playing a digital game (input). The literature distinguishes between positive-affect and negative-affect. The two are mutually exclusive axes, which incorporate several positive and negative emotional dimensions (see Table 2.2).

UXG literature links positive-affect to the constructs of flourishing (which describes well-being, vitality, emotional stability and self-esteem (Russoniello, O'Brien, & Parks, 2009; Vella & Johnson, 2012)), and on a number of occasions flow (Moneta & Csikszentmihalyi, 1996; Nacke & Lindley, 2008; Novak & Hoffman, 1997). Whereas some studies measure the biometric presentation of affect (Mirza-Babaei, Nacke, Gregory, Collins, & Fitzpatrick, 2013), many rely solely on post-game self-report measures. High levels of positive-affect indicate that games can produce positive-affective change that lingers after the game session ends (Ryan et al., 2006), and supports the findings that players use games to effectively manage their mood (Olson, 2010; Olson et al., 2008; Ryan et al., 2006).

2.4.4 Presence and Immersion

Presence is a concept that describes the general sense of "being" somewhere other than where one physically exists (Lombard & Jones, 2015). Researchers often use presence to study how humans experience Virtual Environments (VEs) such as those in digital games. We present a brief overview of presence, and then review its relevance and application to digital games literature. For a more detailed discussion of presence and immersion see Appendix A.

Slater (1999) defines presence by three phenomena:

1. The defining sense of 'being there' in a virtual environment depicted by technology
2. While in a state of presence, humans will respond to virtual events instead of real events
3. After disengaging from the VE, humans have the sense that they have returned from visiting another place

Presence literature presents a long history of redefinition instead of refinement (Lombard & Ditton, 1997; Lombard & Jones, 2007, 2015, Slater, 1999, cf. 2003). Lombard and Jones (2015) provide a thorough review of the formal presence definitions to organise the existing work, and curtail redefinition, misuse, and misappropriations. Some authors (Brown & Cairns, 2004; Jennett et al., 2008; Takatalo, Ihanus, Kaistinen, Nyman, & Häkkinen, 2012) allude to presence because they recognize it as a valid experience in digital games, but generally fail to apply it carefully. For example, *immersion* is a term that is widely used in the digital games community to describe an experience that is both akin to presence while being distinct from it (Cairns et al., 2014). However, the term has specific meaning in presence literature: a situation wherein the sensory stimuli from a virtual environment override those of the real world (Slater, 1999). To further complicate matters, flow literature sometimes uses *immersion* colloquially to describe the essence of the flow experience where one's awareness and attention are consumed, and thus actors are *immersed* in the activity.

There is a large body of research that investigates the concept of immersion as used by the digital games community (Brown & Cairns, 2004; Cairns et al., 2014; Ermi & Mäyrä, 2007; M. C. Green, Brock, & Kaufman, 2004), including the development of a reliable and valid measure (Jennett et al., 2008). This body claims that immersion is multi dimensional, but that some or all dimensions are distinct from existing constructs such as presence and flow. However, their definitions presume presence to be the strongest form of immersion, so-called *total immersion* (Brown & Cairns, 2004; Ermi & Mäyrä, 2007). Other aspects of this immersion, such as engagement, are accounted for in presence literature (Jacobson, 2002; Lombard & Jones, 2015), and the cohort offer little more than anecdotal evidence in arguing that flow is only a rare and extreme experience (Cairns et al., 2014, p. 343–344). Jennett and colleagues (2008, p. 643) argue in a similar fashion and claim that a game cannot provide a flow experience if the goals are unclear, and the feedback delayed. This paradigm presents immersion as a precursor to both flow and presence, and that immersion only becomes flow or presence once the player crosses an unknown and subjective threshold of intensity. Ultimately, the researchers fail to distinguish immersion as a unique experiences. Their conceptualization of immersion is best described as the coincidence of flow and presence at low intesities, but there is no evidence to suggest that it is a unique construct requiring distinction.

The conceptualization and theoretical background for immersion appears tenuous and we do agree that UXG research benefits from considering immersion (Brown & Cairns, 2004) as a unique experience in digital games. Presence is a sophisticated concept that does have relevance in UXG research. The known psychological benefits associated with presence are therapeutic, where it is used as a tool to disrupt patient perspectives, or present patients with safe simulations of distressing situations (Riva et al., 2015). These applications are interesting, but irrelevant to presence in entertainment games. There is little well-researched literature pertaining to the lasting effects on the player when presence is a by-product of gameplay on the player, nor how presence relates to player motivations. If time-limiting policies do influence the experience of presence, there is no theory to suggest that changes in presence degrade the lasting positive effects of playing a specific digital game. In comparison, regular experience of flow produces lasting positive effects and the degradation of flow can influence well-being negatively.

2.4.5 GameFlow

Sweetser and Wyeth (2005) offer one of the earliest considered applications of flow to UXG in their model, *GameFlow*. The authors equate flow and enjoyment calling them “the single most important goal for computer games.” (2005, p. 1544) and proceed to map the elements of flow/enjoyment to elements of usabilty and positive digital game experiences. Their presentation of the term enjoyment is better described by the phrase *positive experience* than by flow. Enjoyment carries connotations of uncomfortable, but meaningful experiences not otherwise addressed by flow (Tamborini et al., 2010). The heuristics presented are derived examples of a relatively new commercial art form. As a descriptive model for UXG, GameFlow describes only replicas of those existing game experiences (K. Poels et al., 2007b) and therefore cannot provide accurate descriptions of novel game experiences. GameFlow is one of the first attempts to describe the holistic player experience, but ultimately stands as a model for succesful game design, rather than a measure of player experience.

Flow itself is a universal and intuitive construct (Csikszentmihalyi, 1990; Engeser & Schiepe-Tiska, 2012). It is therefore easy to use analogy to generate models of how flow manifests in digital games: Chen (2006, 2007) uses the slanty bar analogy and Sweetser and Wyeth (2005) uses existing and succesful games. This approach provides intuitive arguments, but researchers and designers using these operationalizations may overlook important aspects not described by the analogy; only the flow components utilized by the exemplar games form a part of the GameFlow model.

2.4.6 Game Engagement Questionnaire

Brockmyer and colleagues (2009) developed the *Game Engagement Questionnaire* (GengQ) based on established theories of immersion, presence, flow and psychological absorption and dissociation. The questionnaire measures an individual’s tendency to become engaged with number of video games; particularly

Table 2.3 Dimensions of the GEQ Core Module. from (K. Poels et al., 2007b).

Dimension	Experience descriptors
Sensory and imaginative immersion	Story, aesthetics, and engagement
Flow	Concentration, loss of self-consciousness / sense of time
Challenge	Challenge and difficulty
Competence	Skills, ability and performance
Tension/annoyance	Annoyance, irritation, frustration
Negative-affect	Bad mood, bored
Positive-affect	Fun and enjoyment

violent video games. It does not measure the experience of the players in any given session. The measure appears to be valid and reliable, and appears to have potential for use as a state measure of its constructs. Unfortunately, the authors appear to have categorized their items based on how difficult (sic.) they are to experience, not on the theoretical models that informed the questionnaire's creation; for example, they imply that flow is a higher form of engagement than presence, yet all the items associated with presence are components of flow. (See Brockmyer et al., 2009 Fig. 1; and Norman, 2013).

2.4.7 Player experience of needs satisfaction

Self-Determination Theory (SDT; Ryan & Deci, 2000) is a sophisticated model for human motivation that has strong associations with psychological well-being; for these reasons, the theory is associated with flow. Ryan and colleagues (Ryan et al., 2006) draw on SDT's formulation of intrinsic motivation to explain the specific appeal of video games, which is a voluntary[15] activity. They developed the Player Experience of Need Satisfaction (PENS) model, arguing that games easily satisfy an individual's need for *autonomy*, *competency*, and *relatedness,* while also providing an opportunity to experience *presence* and reducing frustration by providing an *intuitive control* mechanism. Autonomy describes a person's volition in an activity, and is supported in games when players are given freedom to behave as they choose. Competence describes a person's need to overcome challenges and have an observable effect on a situation. Relatedness describes a person's need to form social connections with other entities.

The seminal paper (Ryan et al., 2006) uses a long form and short form of the PENS questionnaire that boasts high Cronbach's α scores, but provides only a few of the actual items used to measure each dimension. The full survey does not appear to be readily available. The PENS model is used in other research (Gerling, Birk, Mandryk, & Doucette, 2013; Przybylski, Ryan, & Rigby, 2009) and some publish the items used (Dennie, 2012), but the validity and reliability of those items remain unclear.

[15] Not including laboratory experiments.

2.4.8 Game Experience Questionnaire

Poels and colleagues (2007b) offer a considered model of what they call "the actual experience of playing digital games" (2007b, p. 83), which they later adapted to create the Game Experience Questionnaire (GEQ) (see Table 2.3 p. 27). The GEQ core module measures seven dimensions, many of which relate to the components of flow. *Sensory and imaginative immersion* is the flow dimension *absorption by activity* (Engeser & Rheinberg, 2008) operationalized for digital games. *Positive-affect* is often a consequence of flow (Landhäußer & Keller, 2012). *Challenge* and *Competence* relate to challenge and ability from which one can infer balance. This impressive questionnaire is the best attempt to capture the essence of a generic games experience; however, the GEQ necessarily simpifies its constituent constructs to indentify which factors are present, and their relative intensity in a particular games session. A full report of development and validation is available (K. Poels et al., 2007a, 2007b) although it is not published in academic texts and is difficult to access, with some researchers suggesting it is unavailable (Nordin et al., 2014; Norman, 2013).

2.4.9 Theroretical considerations for this research

It is evident that many researchers have spent many years seeking a unified and succinct means to describe the notion of "games experience". As a result the UXG field appears to be saturated with models and measures based on a variety of theoretical and behavioural theories, developed for sometimes general purposes and sometimes specific purposes. Furthermore, researcher intentions seldom match their position through complexity of the phenomena, and relationships between them. For example, GameFlow (Sweetser & Wyeth, 2005) intends to model experience, but instead provides game design suggestions (Ijsselsteijn et al., 2007). The researchers looking to operationalize immersion as a unique construct do not define it well (Nacke, 2012). Although Jennet and Colleagues' (2008) measure of immersion shows reliability, they fail to present a convincing distinction between their construct and those of flow and presence. Nonetheless, these authors agree that games user experience is too complex to be measured using a single construct (Brown & Cairns, 2004; K. Poels et al., 2007b; Sweetser & Wyeth, 2005; Takatalo et al., 2012).

The instruments themselves can be difficult to access, and very few contain instructions for application, measurment and analysis. The most "complete" instrument appears to be the Game Experience Questionnaire (Ijsselsteijn, de Kort, & Poels, 2013) as it addresses the multi-dimensional nature of UXG for individual game sessions (as opposed to personality traits, or tendencies). It includes only the most sophisticated models of positive experiences commonly associated UXG: namely flow, presence and affect.

Finally, too few of these models provide clear evidence of reliabilty and validity for their construction of UXG. While the instrument may be "good enough", the constituent dimensions possess existing, internationally valid and reliable measures. Therefore, we chose not to use the GEQ, but rather measure flow and affect by using existing valid, internationally reliable and relatively short questionnaires (see Section 3.3.2). We did not perform the psychometric validation ourselves, as this is a non-trivial task, and such an undertaking needs be the subject of a different body of research. Apart from the tension dimension, our choice of measures captures the essence of the GEQ in fewer items.

2.5 Summary and expectations

The literature supports the notion that digital games offer a variety of benefits to players that improve subjective well-being. For example, research suggests that individuals engage with the medium because it helps improve mood, or offers respite from their present life difficulties. The aspect of UXG that is most well-cited to provide this benefit is flow.

The literature further acknowledges that games may be the object of anti-social practice, so-called addiction and sedentary lifestyles, but argues that such incidents are rare, and that digital games do not cause them. It is possible that engagement becomes maladaptive if it distracts sufferers from acting to improve their situation. To combat these negative aspects, policy makers have implemented systems with the specific intention of reducing session lengths. Two of the most prevalent types are the behaviour and design policies. It is unclear how each policy affects session length or those dimensions of UXG associated with well-being.

We expect that flow scores for the behaviour policy will be high; the *cooling off* event may interrupt the player in the throes of a flow experience, especially if the game shuts down without warning. Shutdown events may be especially surprising, given flow's ability to warp the player's experience of time. Flow experiences improve well-being, and we therefore expect to see high positive affect. Furthermore, interruptions to flow are not associated with negative emotions. However, the *cooling off* events are unaware of the player's context during play, thus the policy may frustrate the player's ability to accomplish goals; those they set for themselves in the context of solo play, or their party in the context of group play. We therefore expect to see high negative affect.

We believe that the design policy affects some components of flow. It alters how players can approach character progression, an important goal in the RPG genre. As the fatigued resource becomes rare, the player's perceived ability to overcome perceived challenges should be thrown out of balance. We expect that flow scores will be low, players will become bored and choose to stop playing sooner than they would when not subject to the policies. That being said, if our pilot data is representative, then we expect longer sessions, and will need to find out why. We expect positive affect scores to be high, as the game still provides flow opportunity, and we expected negative to be lower than those seen in the behaviour policy, but higher than unpoliced sessions due to the handicap.

3 METHODS

In the previous chapters, we demonstrated that digital games can have both positive and harmful effects in those who play them because they draw players into compelling experiences. We discussed two distinct policy types that seek to reduce harm done to players by shortening session length either by shutting down the game and ensuring play stops, or by fatiguing game mechanics to discourage continued play. We decided to investigate the utility of such policies in terms of player experience; to see if the policies changed the positive experiences commonly associated with games to negative ones. In this chapter, we describe our chosen methods for addressing these questions. We begin with a summary of our motivating questions and a brief defence of our method choice. The body of the chapter describes our experiment in detail.

3.1 Research questions revisited

See Section 1.3 for a detailed discussion of each question.

- How does the design policy affect session length?
- How do the behaviour and design policies affect player experience in terms of flow and affect?
- How does policy affect a player's intention to return to a game?
- How does the design policy affect a player's reason for choosing to end their session?

3.2 Data collection

3.2.1 Methodology considerations

Large-scale surveys are a common and easy method of data collection for subjective user-experience research in digital games (UXG), particularly when based upon existing systems, as they provide strong external validity. However, we established in

Section 2.2 that the effects of digital games depend upon genre, game, player, and social context (see Section 2.3). This dissertation compares each policy's effects to each other, and to a session representing the absence of policy. To draw conclusions on this, a survey would need to restrict aspects of game, genre, player and context, or collect those variations as additional independent variables. Therefore, we used a controlled laboratory experiment as our method of data collection because it offers the strongest internal validity and, by offers more control over genre, game, player, and context. Thus, we confidently draw conclusions about each policy's effect on players.

3.2.2 Design

Through an experiment we controlled the game and genre variables by making all participants play the same game. We controlled many player and context variables by screening participants, maintaining a consistent schedule and environment, and controlling the social context in which they played. To help control for player factors we used a *repeated measures* design that borrows many elements from a *cross-over* experiment design. A common design for clinical trials, the cross over design seeks to identify and examine differences between a number of treatments, including a placebo or control treatments (Senn, 1993). Participants experience each condition at least once to represent each exposure. The design must vary the treatment sequence and must include a 'washout' period between each treatment to minimize carry-over effects and interaction effects. The cross-over design offers more control than a between-subjects design because the effects of participant variables on the measured outcome are minimized with each individual serving as their own control (Girden, 1992). This was important for our research because of the potential influence of player factors upon the effects (see Section 2.3.2, and discussion and conclusion of between-subjects pilot in Appendix E). For example, our cross-over pilot study (Appendix H) showed large differences in affect scores between subjects playing under the same conditions. This result indicates that different players have different experiences, or different intensities of experiences, based on their own individual characteristics.

3.2.3 Independent variables

Our research seeks to identify the effects of policy on player experience and session length. Therefore, *policy* acted as our independent variable and had three levels, or experimental conditions: *shutdown*, *fatigue,* and *control*. The *shutdown* condition simulated the essence of the behaviour policies described in Section 1.2.1 by forcing players to end their session after a specified period. The *fatigue* condition represented the design policies as described in Section 1.2.2. Players experienced unmodified gameplay for a period, after which character progression rates and loot-drop rates tended towards zero. The remaining level, *control,* represented the absence of a specific policy; other than those intended by the game designers, the condition did not manipulate game mechanics, did not impose a session length upon players. This condition is sufficiently different from the shutdown and fatigue conditions to act as a baseline against which we compared results.

We identified in Section 1.2 that the cooling off policy allows two hours of play before forcing hiatus, then allows an additional hour. The GAPP requires the fatigue system begin after a three-hour period, and peak after two further hours. Our pilot data (based on 10 participants) suggests that participants will play for an average of two hours under laboratory conditions Appendix H. For this reason, we must contrive to ensure participants experience the shutdown and fatigue events so that we may observe their reactions. Thus, the events were set to occur after one hour of play.

3.2.4 Dependant variables

Our measured outcomes include *session length*, *flow* and *affect* dimensions of UXG, the player's *intention to return* to the game, and *reason for hiatus*. A hiatus is any pause or break in a sequence or activity, therefore reason for hiatus is a reason why players chose to stop.

The behaviour and design policies described in Section 1.2 are attempts at managing, or reducing, session length. Therefore, we measured the length of all play session that occur in the experiment.

See Section 2.4 for detailed discussion of player experience. We support the argument that player experience is multi-dimensional, incorporating aspects of flow and affect. Furthermore, we support the argument that the effects of games result from player experiences, which are themselves functions of game, genre, player and context. We operationally define player experience as a multi-dimensional concept that includes flow and affect. An understanding of how policy affects flow can grant understanding of how long-term exposure to policy can influence the utility provided by flow (see Section 2.4.1). An understanding of affective responses to games allows us to make inferences about the potential for positive emotional experiences as well as emotion regulation (see Sections 2.2.1 and 2.4.3). We did not include presence in our operational definition because we found no evidence of lasting effects in player well-being.

3.2.5 Other considerations

Williams (as cited in Cochran & Cox, 1957) argue that carry-over effects are inherent in cross-over studies and are difficult to control for, and often should not be ignored once control mechanisms are in place. They reason that these effects are relevant to explaining the observed results of the dependent variable, and rather than being controlled for, researchers should identify them. Furthermore, while participants act as a control unto themselves we cannot control their effect on context: such as day-to-day changes that require different behaviours. Examples may include once-off commitments and momentary emotional stress. Through observation of the participants, we can better understand the influence of carry-over effects. To this end, we asked participants to comment on their decision to end their session. Furthermore, we recorded conversations and observable behaviours during each session from which we extracted qualitative information.

Our experiment made use of human subjects. We attained ethical clearance from the Faculty of Science Research Ethics Committee at the University of Cape Town (see Appendix F).

3.2.6 Hypotheses

Our experiment aimed to test the following hypotheses:

a. The type of policy has an effect on game sessions-length.
b. The type of policy has an effect on the experience of flow state.
c. The type of policy has an effect on Positive-affect.
d. The type of policy has an effect on Negative-affect.
e. The type of policy has an effect on intended hiatus-length.

3.3 Materials

3.3.1 Experimental system

We modified the critically acclaimed Torchlight II (Runic Games, 2012) to employ both design and behaviour policies. In doing so we maximised the internal validity of the experiment by ensuring that players experienced enough engaging content to at least be engaged for the duration of the experiment.

Torchlight II is an action Role-Playing Game, a genre with mechanics often studied in problematic gaming research (Griffiths, 2016). In Torchlight II, players create a *character*[16] by selecting a *class*[17] and custom appearance that becomes their *avatar* within the game world (see Figure 3.1 p. 35). Players then name their character and choose, from a limited pool of animals, a pet that assists them in their journey. Players view the world and their character from an isometric perspective and are free to explore the game world of wilderness and dungeon, killing monsters, completing quests, and searching for treasures in a rich fantasy world (see Figure 3.2 p. 36). If a player's character is killed, they can *respawn*[18] in a safe area, such as a town, where they cannot be attacked, or closer to the area where they were killed for some cost to their progress.

[16] An avatar for themselves or some champion whom they wish to role-play. Players play the game by controlling this avatar.

[17] Class defines a character's unique strengths, weaknesses, and skill-set.

[18] Their character is resurrected in the game world.

Figure 3.1 The *Torchlight II* character creation screen. Players can select a class (outlined), customize their appearance (outlined), choose a name (outlined) and select a pet; here a deer.

Table 3.1 Values governing the fatiguing of loot. Values chosen so that loot rates reach (effectively) zero at between 63 and 65 minutes.

Fatigued stat	Effect Name (as per GUTS)	Amount
XP Gain	XP GAIN BONUS	-5%
Magical drop	PERCENT MAGICAL DROP	-20%
Gold found	PERCENT GOLD DROP	-10%
Gold drop	GOLD DROP	-10%

Character building is a primary objective for players in Torchlight II is. Experience Points (XP) are the main resource that allows players to build their characters. Players collect XP whenever they defeat enemies. Whenever the amount of XP collected exceeds a threshold, the player's character gains a level (see Figure 3.3 p. 36). Whenever players gain one level's worth of XP they can increase, by a limited amount, their character attributes: Strength, Dexterity, Vitality and Focus[19]. Their character can also learn one new skill or spell, or improve one learned skill. All other mechanics that contribute to character progression are generally known as *loot*. Loot includes items such as weapons, armour and other treasures, and gold. Items have attributes that increase character abilities, making them stronger. Players use gold to purchase items for their character at shops, and to respawn their character in convenient locations. For the purposes of this research, we included XP as a form of loot because it serves the same purpose as the other types of loot: to make the character more survivable. Players acquire Torchlight II XP in the same manner as other loot: defeating enemies. Loot is plentiful in the unmodified version of Torchlight II, but as the player's progress through the game and their characters become stronger loot becomes increasingly better and more interesting, but is less likely to drop.

[19] These govern mechanics such as weapon damage, chance-to-dodge in-coming attacks, maximum health, and spell damage respectively.

Figure 3.2 Gameplay in Torchlight II. The image shows a player's character fighting a monster. (outlined). At the bottom of the screen is information for the player about their character. The blue bar at the very bottom is an experience bar. It gives an indication of how much experience the player has – currently 45% of their current level requirement. When the bar fills up, the character levels up. See Figure 3.3.

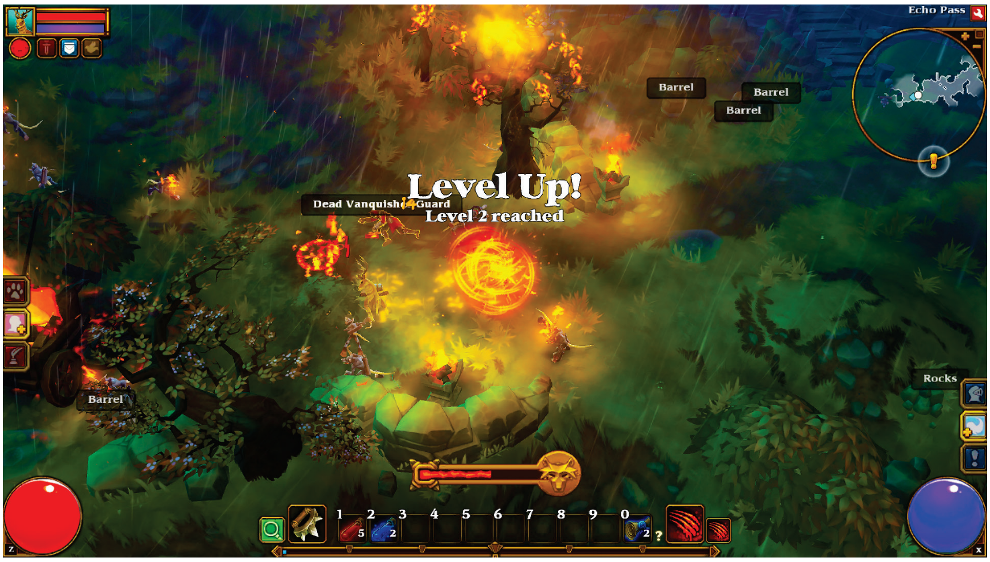

Figure 3.3 Character level up. The experience bar was filled, and so the player's character levels up, and the experience bar empties. Also seen on screen is other loot, dropped by slain enemies.

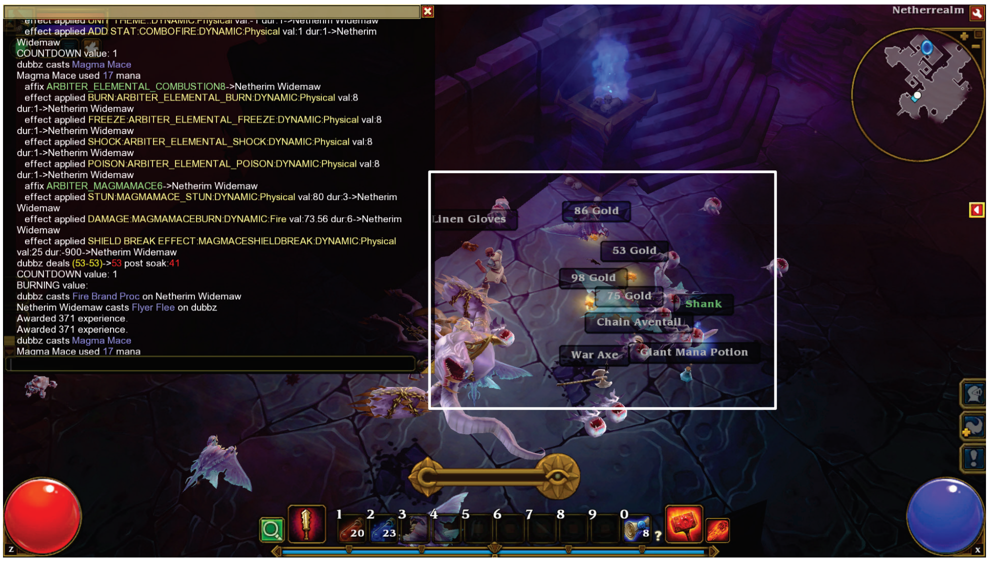

Figure 3.4 Normal loot drops. The figure shows a player character unaffected by the fatigue mechanics. The smaller text boxes (outlined) show that loot has dropped on to the floor. The amount shown is normal for this dungeon. White text indicates items are of normal rarity, whereas green text the item is rarer, and has magical properties. The text box in the top left corner shows the combat log including experience awarded. See figure below for combat log.

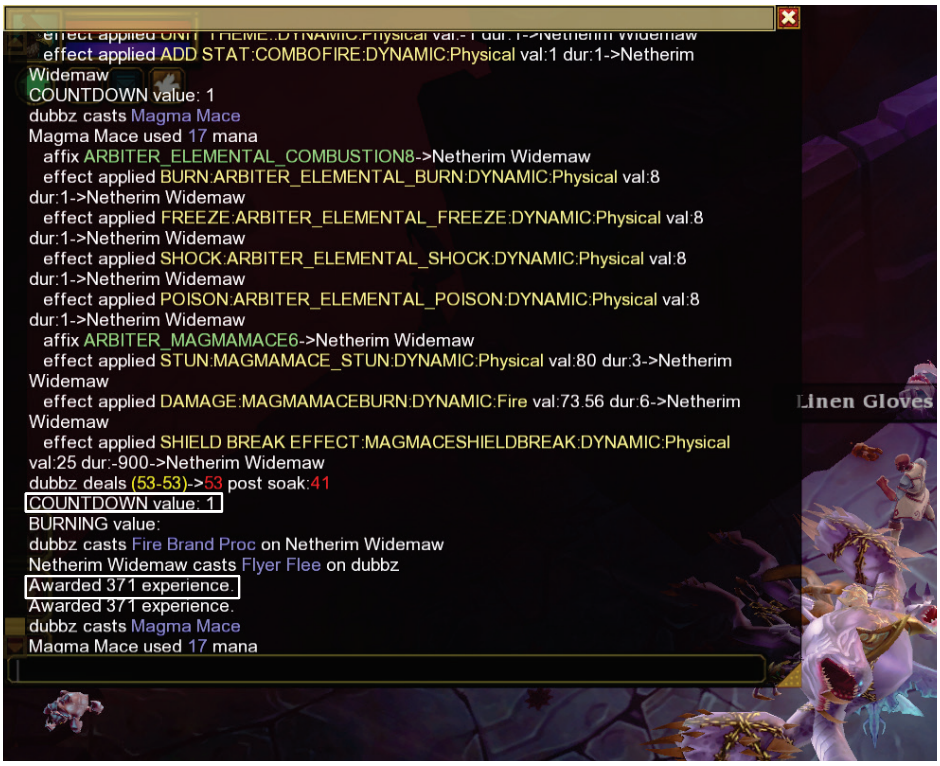

Figure 3.5 Combat log for control condition. Here, the COUNTDOWN value (outlined) is the positive integer 1. The fatigue mechanics function only if the COUNTDOWN value is zero. The game also awarded a normal amount for experience (outlined) for this dungeon.

Figure 3.6 Fatigue loot drops. The figure shows a player character that is affected by the fatigue mechanics in the same room as Figure 3.4 (p. 37). Text boxes (outlined) show loot dropped on the floor on the floor. There are no items, and only three gold piles of 1 Gold on the floor, compared with the 53-98 Gold in Figure 3.4 (p. 37). The text box in the top left corner shows the combat log. See figure below.

Figure 3.7 Combat log for fatigue condition. The COUNTDOWN value (outlined) is zero, thus the fatigue mechanics influence the game. This excerpt shows the RUNN_PERCENTGOLDDROP_FATIGUE (outlined) mechanic at work. The mechanic reduces gold drop by a value of 10 for each instance of damage dealt by the player. The excerpt also shows that zero experience is awarded (outlined), and that the Gold Find value is negative 108 (outlined).

Modification

The developers of Torchlight II released a content creation tool, GUTS (Runic Games, 2013) that allowed us to modify Torchlight II and introduce to the game a set of loot fatigue mechanics to the such as those seen in behaviour policies. Figure 3.4 (p. 37) and Figure 3.5 (p. 37) show the types and quantity of loot that drop in normal gameplay. We created a modification that used a countdown timer set 3700 seconds. After the timer reached zero, the variables governing loot drop-probability decreased by a percentage with every instance of damage dealt by the player's character (see Table 3.1 p. 35). The modification engine did not register any changes less than 1%, thus our design policy does not account for an extended period of reduced loot. Rather, the variables very quickly tended towards zero, severely stunting the rate at which players acquired fresh loot (see Figure 3.6 (p. 38) and Figure 3.7 (p. 38) for an example of how this looked).

The shutdown system required that participants stop playing abruptly and absolutely after a 60-minute period. We chose this based on our pilot data (see Appendix H), where players spent an average of 90 minutes in the unmodified version of Torchlight II. We did this to increase the probability that participants experienced the shutdown event before they felt ready to go on hiatus. Torchlight II and GUTS do not provide a method to implement a shutdown event from within the game. A Microsoft Batch script (Figure 8.1 p. 116) sufficed to simulate the policy. The script launched Torchlight II and started a timer counting down from 3700, after which it executed a system call to Windows, saved the player's progress and killed Torchlight II's process.

The control condition required no modification. Players merely played the unmodified version of Torchlight II.

3.3.2 Instruments

We used a clock to measure session length and rounded up to the nearest minute. Higher levels of accuracy for this session length are inconsequential.

We used The Flow Short Scale (FSS) (Rheinberg, Vollmeyer & Engeser, 2003; cf. Engeser & Rheinberg, 2008) to measure flow. The instrument measures the levels of perceived absorption and performance, which together make up a measure of flow experience. In addition, the instrument measures how important the activity feels to participants. Finally, it measures the perceived levels of challenge, ability, and the perceived balance of those two factors. The FSS is internationally validated and shown to be reliable, where the authors report Cronbach's $\alpha = 0.9$ (Rheinberg, Vollmeyer, & Engeser, 2003). The scale has seen some use when studying games (Engeser & Rheinberg, 2008; Weibel & Wissmath, 2011; Weibel, Wissmath, Habegger, Steiner, & Groner, 2008)

The International Positive and Negative-affect Schedule Short Form (I-PANAS-SF) (Thompson, 2007) was used to measure affect. Much like the PANAS (D. Watson et al., 1988), the I-PANAS-SF has been rigorously validated, and found to be reliable for a wide range of nationalities. Not only does the short form of the schedule require half the time to administer, it is also more appropriate than other affect measures for a wider range of cultures as found in many student populations. The instrument uses Likert-type scales to measure the extent to which respondents feel positive: *alert*, *inspired*, *determined*, *attentive* and *active*. It further measures the extent to which they feel negative: *upset*, *hostile*, *ashamed*, *nervous* and *afraid*. Although the instrument was developed as a trait measure, it has been used with success as a state measure a number of times (Falconer et al., 2014; Fredenburg & Silverman, 2014; Karim, Weisz, & Rehman, 2011; Kuesten, Chopra, Bi, & Meiselman, 2014; Sanchez, Moss, Twist, & Karageorghis, 2014) in games research on occasion (Vermeulen, Castellar, Janssen, Calvi, & Van Looy, 2016).

We collected participant's experiment data in LimeSurvey (LimeSurvey GmbH, 2003), an online survey and questionnaire tool. Questionnaires were transposed into LimeSurvey forms.

3.4 Participants

3.4.1 Sample size

The cross-over nature of the design required that each member of the sample is subjected to all interventions, and were tested after each intervention. Three separate interventions required a significant time investment from the participants. Our design saw participants begin each session at the same time of the same day each week. This meant each participant had a one-week[20] washout period and three-week turnover (see Section 3.2.2). As the number of sessions increased, so did the attrition rate. Failure to complete all sessions would mean that a participant's data could not be used for analysis, as there is less control over player variables in these instances. Thus, we provided a monetary incentive of one hundred and twenty South African Rands due to those who completed the experiment.

A *Repeated Measures ANOVA* is the most common inferential statistic test used for cross-over studies. We used G*Power's (Faul, Erdfelder, Lang, & Buchner, 2007) *a-priori* power analysis tool to estimate the sample sizes required for adequate statistical power. Using SPSS 23 (IBM, 2014) and data from pilot studies (see appendices D and E); the variables required for the sample size calculation are estimated (see p. 41). We use the SPSS calculation for partial eta-squared (η_p^2) to describe effect size here.

[20] Minus their previous session's length

Table 3.2 Values used for calculating required sample size in G*power. We used the F tests: ANOVA: Repeated measures between factors. Values for each dependent variable estimated from pilot-study data. Error probability α is 0.05 and power is 0.95 for all DV (For FSS, we also used the lower bound of generally acceptable power of 0.8 (Cohen, 1988) to identify smallest acceptable sample size. Number of groups and Number of measurements both are 3. Options: Effect size specification as per SPSS. For a complete discussion, see Appendix H.5.

			Effect Size		Required
Dependent Variable	α	Power	SPSS η_p^2	f(U)[21]	Sample Size
PA	.05	0.95	0.294	0.645	45
NA	.05	0.95	0.420	0.851	30
FSS	0.5	0.95	0.113	0.357	129
	.05	0.8	0.113	0.357	84
Session length	.05	0.95	0.794	1.96	12
• Control & Fatigue	.05	0.95	0.356	1.00	36

3.4.2 Population considerations

Our sample must include only those with sufficient computer and games experiences to minimize the effect that novelty might have on measured outcomes such as flow[22]. We recruited participants from the University of Cape Town (UCT) student population who play games. Additionally, all estimates of population include all gamers: those who play on their personal computers, mobile devices and consoles. Mainstream PC games presume that players are proficient users of the keyboard and mouse, whereas console gamers are proficient with archetypal game controllers. There is no guarantee that the entire gaming population are familiar with every metaphor. Our choice of gaming platform will have consequences for our choice of participant.

In South Africa, the university demographic grew up in a time of limited computer access. Hence gaming was a luxury afforded to only the privileged. In 2004, an estimated 12% - 13.6% of households owned a PC, with provincial percentages varying from 4.4% in Limpopo to 33.8% in the Western Cape (Thlabela, Roodt, Paterson, & Weir-Smith, 2006). UCT accepts students from many different provinces and nations, with varied levels of exposure to computers. Differing levels of familiarity with computers is a player variable not accounted for by the cross-over design. Thus, we screen each applicant to ascertain their experience with computers and games and exclude those where the novelty of the gaming experience may offset the effects of the policy. South African law mandates that for an individual to make a properly informed decision, they must be at least 18 years of age. Thus, it was required that all participants be 18 years or older.

[21] Calculated by G*Power

[22] For example, unfamiliarity with the game's interface may contribute to flow's challenge dimension more than the game itself.

Table 3.3 Policy-exposure order. Show the three participant groups, and the sequence in which they experienced each policy.

Participant group	First session	Second session	Third session
Sequence *SCF*	Shutdown	Control	Fatigue
Sequence *FSC*	Fatigue	Shutdown	Control
Sequence *CFS*	Control	Fatigue	Shutdown

3.4.3 Recruitment

The experiment was advertised around the campus of UCT. All participants agreed to take part in the experiment of their own free will and indicated so on a consent form (see Appendix D) signed by both them and the present experimenter. The consent form also served to collect demographic information relating to the participant's computing and gaming experience. This information served as a screening process to exclude those people who might skew the data if their experience reflects the act of engaging with a novel object, rather than the act of playing a game. Participants indicating less than a year's experience with computing, who had never played more than an hour of gaming per week for an extended period, had not played games on a PC before and had no experience with RPGs or similar genres were thanked for their interest, but turned away from the experiment. Participants were aware that those who completed the study received money (ZAR 120) both as compensation for assisting with the research, and as incentive for returning to complete the experiment.

A total of 33 gamers between the ages of 18 and 39 (In years: M = 22.78, SD = 3.83; 12% female and 88% male) participated in the experiment. Twenty-nine were students, twelve of whom were studying computer science with four of those twelve taking courses in game development. All participants indicated that their experience using a computer was of at least a moderate level (Number: "Moderate" = 10, "Lots" = 17, "Master" = 6). Participant's gaming behaviour at the time of the study varied, with hours per week ranging between zero and 100 (In hours: M = 13.34, Median = 7.5, SD = 20.04, Min = 0, Max = 100). Thirty indicated that at the time of the study, they played games for at least one hour per week. Those three who indicated otherwise had played games on a regular basis in the past of at least three hours per week. Twenty-eight participants indicated prior experience with Torchlight II itself, or an isometric action RPG with similar mechanics. The remaining five had played games that fall under the super-genre, so called RPG, that contain similar mechanics and user interface (In hours: M = 117.96, SD = 148, Non-zero min = 5, Max = 500).

3.5 Procedure

Participants were required to attend three play sessions and were randomly assigned to one of three balanced participant groups. Each group experienced the three conditions in a unique sequence. We used a Latin square method to vary the order of treatment and counterbalance the order of exposure to help minimize and identify order effects (see Table 3.3; see Section 6.3 for discussion of shortcomings).

For the first session only, before it began in earnest, each participant read through a document informing them of their role, requirements, anonymity, rights and recourse for the experiment. The researcher and participant both signed two copies of the document; each party kept a copy for themselves.

Participants were then told which policy they would experience during their impending session. Leading into the shutdown condition, players were informed that after an hour of normal play, the game would save their progress and exit and that they would not be allowed to continue. Before the fatigue sessions, the experimenter informed players that the game would begin to fatigue after an hour of normal play, and identified the relevant mechanics. Participants in the control condition were told that could play for at least an hour, but could continue playing for as long as desired. Neither behaviour nor design policy described in Section 1.2 are secret implementations; the policies are intended to be a permanent addition to gaming practice and players know they exist. That knowledge certainly affects the decisions players make, and the behaviours they take. Full disclosure of the conditions is imperative to the ecologic validity of the experiment because. Thus, players subjected to these policies would be expecting the events. We discuss implications for this decision in Section 6.3.

In their first session, participants were free to select a class, and create a character to play with, or as, for the duration of the experiment. See Section 3.3.1 for character creation details. Once the session had come to an end, the participants were presented with the LimeSurvey website. Here, they completed the affect schedule to indicate how they were feeling at that specific point in time. They were then presented with the flow questionnaire and asked to complete it in terms of how they were feeling for the final minutes prior to the point at which they went on hiatus. Participants were finally asked to provide additional information to help qualify the nature of their hiatus behaviour. First, they were asked to explain why they had stopped playing. Then, they were asked to select from a list of options where in the game world their character was located when they stopped playing ("In town/camp", "In the wilderness", "In a dungeon"). This information might describe interesting behaviour such as preparing their character for the hiatus. Players were asked to comment on their goal or purpose in that area, and the in-game activity that their character was pursuing in that area (fighting, fishing, exploring, questing, etc.) Finally, players were asked to indicate their current desire to continue playing the game, and to comment on why they felt this way. This information would describe participant's intentions to return.

After the final session, participants were thanked for their help, and given the opportunity to ask any questions, or voice any concerns that they might have had. If any participants had felt distressed by the experiment for any reason at any stage and wished to desist, they would have been allowed to and advised to seek assistance from the university's student health services or would have been referred to a clinical psychologist. We downloaded all participant data from our survey hosting LimeSurvey and imported it to IBM SPSS 23 for analysis

4 RESULTS

4.1 Introduction and statistical methods

In this chapter we analyse the data we collected during our experiment. We begin with a high-level summary of the main findings. We then present descriptive statistics and inferential analyses in terms of p-values and effect sizes for each construct measured in the experiment. Our experiments consisted of three conditions: The fatigue condition wherein the rate of character progression slowed after an hour of play; the shutdown condition wherein the system quit after an hour of play, and the control condition wherein the system allows normal play.

To identify effects between conditions we chose the most appropriate statistical tests for our data set: *repeated measures ANOVA,* test unless the data violates the assumptions of normality, in which case we used a *Friedman's non-parametric* test with a post-hoc *Wilcoxon sign-rank* test. See Table 4.1 p. 46.

Our experiment uses a cross-over design to help control for carry-over and sampling effects (see Section 3.2.2). We acknowledge that these effects are always present in experiments (see Section 3.2.5). We therefore seek to understand the influence of carry-over effects on our results by grouping our data according to their ordinal value. This allowed us to compare all the sessions that occurred first, with all the sessions that occurred second, and all the session that occurred third. We seek to understand the influence of sampling effects on our data by testing for interaction effects, and by grouping our data according to their participant group where appropriate. For example, we group all data points that experienced the sequence shutdown followed by control and then by fatigue (see Table 3.3 p. 42). We then perform a repeated measures ANOVA or Friedman's test for these groups.

Table 4.1 Summary of statistical methods used and results.

Factor	Significant result (dimensions)	Effect size[23] (absolute difference)	Conditions involved
Session length	Yes	Large (23 minutes; 43 minutes respectively)	Fatigue > Control > Shutdown
Flow	No (0/6)	Inconsequential	None
Positive-affect	No (0/5)	Inconsequential	None
Negative-affect	Yes (1/5)	Small (1 out of 25)	Shutdown > Control
- Upset dimension	Yes	Small (1 out of 5)	Shutdown > Control
Intention to return	Yes	Small (1 out of 6)	Shutdown > Fatigue
Reason for hiatus	Yes	NA	Fatigue and Control

4.1.1 Summary of results

The fatigue condition produced a significant effect of moderate size for the time factor (Section 4.2) where participants played for *longer* in the fatigue condition than in the control-condition. The shutdown condition produced significant but small effects for negative-affect (Section 4.4) and intentions to return (Section 4.5) factors, where scores are higher than the control-condition and fatigue condition respectively. We found no significant effects for flow (Section 4.3.1) and positive-affect (Section 4.4.1).

4.2 Session length

Over the course of three sessions, participants played the experimental system (Torchlight II) for at least 207 minutes, at most 432 minutes and an average of, M = 291 minutes with a standard deviation of 50 minutes. All values in this section are in minutes. We hypothesised that conditions would create an effect for the length of time players spend in game. All conditions show mean differences (see Table 4.2 on p. 47), where the fatigue condition presented the longest average length of 127 minutes (see Figure 4.1 on p. 47). The shutdown condition's period is shortest at a fixed 60 minutes. The control condition's period falls between the other two with 103 minutes. Participants 32 and 33 recorded fatigue session lengths of 242 minutes and 225 minutes respectively contributing to the conditions high mean of 145 minutes. Both qualify as outliers because they exceed the upper limit of normal data. Their exclusion does not significantly change the result of our hypothesis test, nor do they affect our inferences onto the population. Thus, we include their results for all further tests. See Table 4.3 (p. 48).

[23] Based on suggestions given by Cohen (1988) and Miles & Shevlin (2001) as cited in (P. Watson, 2011)

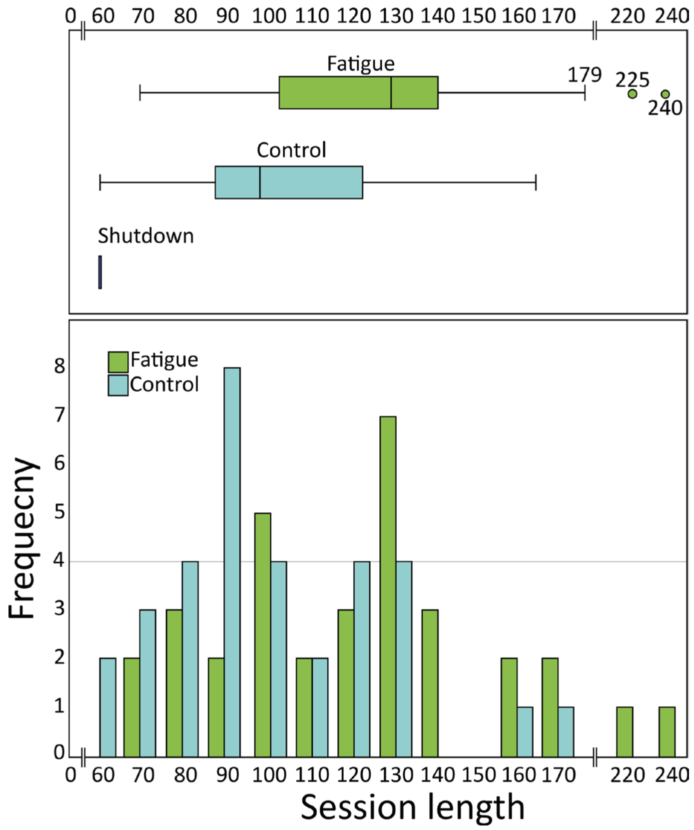

Figure 4.1 Frequency distribution and Box & Whisker of session length in minutes.Shutdown condition is fixed at one hour. Participants 32 and 33 recorded fatigue times that qualify as outliers. Discussion of their specific cases occurs in Section 4.2. Session lengths in the control tend to fall in the shorter categories, and the IQR occupies a lower space in the domain.

Table 4.2 Descriptive stats for session length. All values in minutes, except for column 'N' which is in units. Total based on all 99 data points, not condition values presented in table.

					95% CI for Mean		
Session	N	Min	Max	Mean	Lower	Upper	SD
Fatigue	33	70	242	127	114	141	39
Control	33	60	290	103	94	113	26
Shutdown	33	60	60	60	60	60	0
Total	99	207	432	291	273	309	50

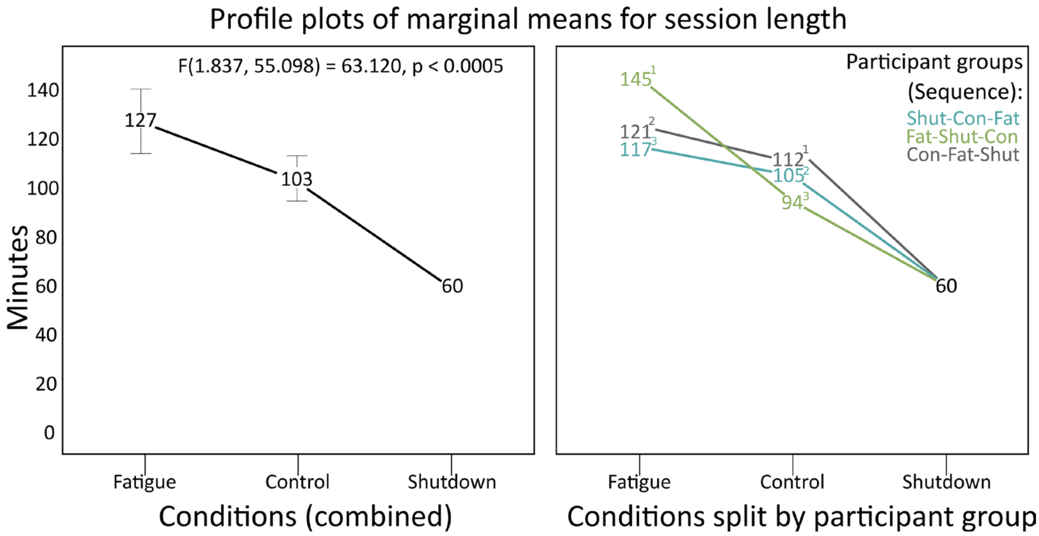

Figure 4.2 Profile plots describing the effects found in the repeated measures ANOVA test of session length.

Numbers indicate the estimated marginal-mean values. The combined graph does not differentiate between participant groups. Whiskers indicate 95% confidence interval for the mean. Colour indicates the participant group (see Table 3.3) from which the data originates. Superscript numbers explicate the ordinal value of the condition for that group. For example, the blue graph represents the group of participants that played shutdown first, control second, and fatigue third. The high Fat- Shut-Con fatigue mean of 145 minutes is discussed in Section 4.2. See Figure 4.1 (p. 47) and Table 4.2 (p. 47) for descriptions.

Table 4.3 Inferential stats for session length.

Shows statistics for data with outliers included (N = 99) and removed (N = 93). No consequential changes between both sample sizes.

Outliers	N	F	df	Sig. (p)	η_p^2	95% CI for η^2 lower	95% CI for η^2 upper	power
Included	99	63.12	1.837	< 0.0005	0.678	0.518	0.759	1.00
Removed	93	65.684	2	<0.0005	0.701	0.527	0.763	1.00

4.2.1 Main effects

We test our hypothesis using a repeated measures ANOVA test. Thus, we hypothesize that H_1: Policy has an effect on session length, or H_0: policy has no effect on session length. The data meets all of ANOVA's assumptions except for sphericity, where we used the Huynh-Feldt correction for degrees of freedom. The test identifies a significant effect between factors, $F(1.837, 55.098) = 63.12$, $p < 0.0005$, $\alpha = 0.05$. Thus, we accept H_1 and see that the effect is large, $\eta_p^2 = 0.678$, see Figure 4.2 and Table 4.3 for further details.

All three conditions are significantly different from each other. The average difference between the fatigue and control condition is 24 minutes ($p = 0.005$, $\alpha =$

0.05, 95% CI [6, 41])[24]. The remaining pairs both have significant effects with $p < 0.0005$ with a mean difference of at least 40 minutes (see Table 4.4 p. 49 for details).

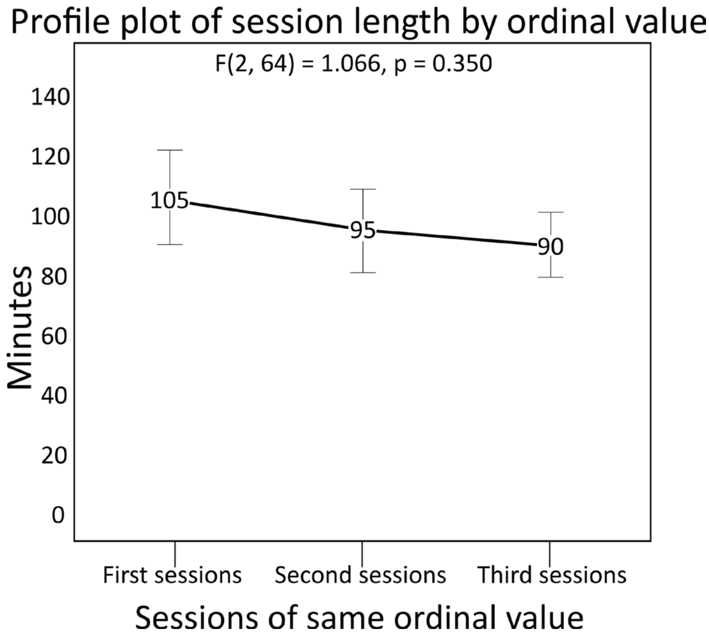

Figure 4.3 Profile plot for repeated measures ANOVA test run between sessions grouped by their ordinal value.Numbers indicate the estimated marginal-mean values, whiskers indicate 95% confidence interval for the mean.

Table 4.4 Descriptive stats for time spent playing, sessions grouped by their ordinal value. Total based on all 99 data points, not condition values presented in table.

					95% CI for Mean		
Session	N	Min	Max	Mean	Lower	Upper	SD
First	33	60	242	105	88	122	47
Second	33	60	179	95	82	108	36
Third	33	60	142	90	79	101	30
Total	99	207	432	291	273	309	50

The shutdown condition has a fixed period, meaning that the value does not vary at all. Whereas SPSS's ANOVA provides a test for significance on each condition pair, it does not provide an effect size. We therefore perform a paired samples T-test to compare the fatigue and control conditions alone. There is a significant difference in session length $t(32) = 3.188$, $p = 0.003$, $\alpha = 0.05$. We accept H_1 and observe a moderate effect size of $d = 0.57$. Here we calculated effect size Cohen's d using Wiseheart's (2013) online calculator for paired samples t-tests, which uses Morris and DeShon's formula 8 (2002) takes into account the correlation between means $\rho = 0.168$.

[24] Adjusted for multiple comparisons, using Bonferroni

4.2.2 Carry-over and sampling effects

No significant effects are present between our subjects, $F(2, 30) = 0.259$, $p = 0.773$, $\alpha = 0.05$, suggesting that sampling effects are inconsequential. Furthermore, removing the outliers from the dataset does not significantly change our results. We are confident that our sample adequately represents the population.

We observe absolute differences between the data when grouped by their ordinal values (see Figure 4.3). There is a mean difference of 10 minutes between the first sessions and second sessions, and a mean difference of 15 minutes between the first and third sessions (see Table 4.4 p. 49). This difference appears large, but we do not find any statistical evidence suggesting this difference is consistently observed in the population, $F(2, 64) = 1.066$, $p = 0.35$, $\alpha = 0.05$. Furthermore, the 95% confidence intervals suggest that the absolute differences are inflated due to participants 32 and 33 who both submitted their outlier session firsts. These differences are inconsequential as would not likely be present in the population. Thus, we are confident that our experimental condition causes the main-effects and are not significantly influenced by the carry-over and sampling effects.

Table 4.5 Descriptive statistics for the Flow Short Scale's core factors. Shows flow and its constituent factors Fluency of performance, and Absorption by activity. Scores are on a scale of 1-7. The midpoint of the scale is four. Totals based on all 99 data points, not condition values presented in table. Cronbach's α = 0.84 based on our data.

						95% CI for Mean		
Factor (items)	Condition	Number	Min	Max	Mean	Lower	Upper	SD
Flow (10)	Fatigue	33	2.9	6.6	5.21	4.76	5.46	0.99
	Control	33	3.2	6.7	5.54	5.13	5.84	0.99
	Shutdown	33	3.8	7	5.35	5	5.57	0.81
	Total	99	2.9	7	5.29	5.10	5.48	0.93
Fluency of performance (6)	Fatigue	33	2.83	7	5.38	4.98	5.76	1.1
	Control	33	3.67	7	5.71	5.37	6.06	0.99
	Shutdown	33	3.5	7	5.5	5.18	5.82	0.89
	Total	99	2.83	7	5.53	5.33	5.73	1
Absorption by activity (4)	Fatigue	33	2.5	6.75	5	4.29	5.16	1.22
	Control	33	2.5	6.75	5.5	4.72	5.56	1.18
	Shutdown	33	3	7	5	4.61	5.33	1.01
	Total	99	2.5	7	4.95	4.71	5.17	1.15
Importance (3)	Fatigue	33	1	7	3.52	2.92	4.12	1.71
	Control	33	1	6	3.32	2.78	3.87	1.53
	Shutdown	33	1	7	3.76	3.21	4.31	1.55
	Total	99	1	7	3.53	3.21	3.85	1.59

4.3 Flow

We measured flow state using the *Flow Short Scale* (FSS) (see Section 3.3.2 and Appendix C). The scale has three main sections, namely *flow* (which is split into dimensions *fluency of performance* and *absorption by activity), importance*, and *challenge/skill*. The FSS uses a scale of *1-7* for the items measuring both dimensions of flow and for the items measuring importance (see Table 4.5). The items for the remaining section are measured on scale of 1-9. Final scores for each section (see Table 4.6 p. 51) are the average of its item scores[25].

Overall, participants appear to experience a moderate to high intensity flow state towards the end of each session. For flow and its dimensions, the control condition produced the highest mean and Confidence Interval (CI), and the fatigue condition produced the lowest mean and CI. See Table 4.5

Importance describes how much consequence a person assigns to the outcome of the activity, or how much of an impact the outcome will have on their life. Participants reported mean values of around M = 3.5. The shutdown condition produced the highest mean and CI, while the control condition produced the lowest. See Table 4.5

Participants perceived that Torchlight II was a low demand activity, for which they themselves possessed moderate to high skill levels. They indicated that the demand was "just right" for both the fatigue and control conditions (M = 5), while the demand during shutdown condition was slightly too easy (M = 4). See Table 4.6 p. 51

Table 4.6 Descriptive statistics for the Flow Short Scale's Challenge and Skills factors. Each item is a separate construct. Scores are on a scale of 1-9. The midpoint of the scale is five. Totals based on all 99 data points, not condition values presented in table. Cannot calculate Cronbach's α for individual items.

Factor (items)	Condition	Number	Min	Max	Mean	95% CI for Mean Lower	95% CI for Mean Upper	SD
Demand (1)	Fatigue	33	1	9	3.39	2.64	4.15	2.14
	Control	33	1	7	2.85	2.34	3.36	1.44
	Shutdown	33	1	9	3.33	2.66	4.01	1.90
	Total	99	1	9	3.19	2.82	3.56	1.84
Skills (1)	Fatigue	33	3	9	6	5.45	6.86	1.98
	Control	33	2	9	7	5.89	7.19	1.82
	Shutdown	33	2	9	7	5.73	7.06	1.89
	Total	99	2	9	6.36	5.98	6.74	1.89
Perceived fit of demand and skills (1)	Fatigue	33	2	8	5	4.19	5.08	1.25
	Control	33	2	6	5	4.02	4.7	0.96
	Shutdown	33	2	7	4	3.76	4.73	1.37
	Total	99	2	7	4.41	4.17	4.65	1.2

[25] As per convention in the literature.

4.3.1 Main, carry-over, and sampling effects

For all factors in the FSS, we test our hypothesis using a repeated measures ANOVA test. Thus, we hypothesize that H_1: Policy has an effect on flow state, or H_0: policy has no effect on flow state. The data meets all of ANOVA's assumptions except for sphericity, where we used the Huynh-Feldt correction for degrees of freedom. We found no significant main effects for condition in any flow factor for $\alpha = 0.05$[26] (see Table 4.7), nor did we find interaction effects based on participant's group or ordinal value for $\alpha = 0.05$. Thus, we fail to reject the null hypothesis that these policies have no effect on flow as measured by the FSS. See Appendix H.3 for details of all test results.

Table 4.7 Inferential statistics for the Flow Short Scale.

					95% CI for η_p^2		
Component	F	df	Sig. (p)	η_p^2	lower	Upper	Power
Flow	2.95	2	0.060	0.090	< 0.0005	0.225	0.554
Importance	1.69	2	0.193	0.053	< 0.0005	0.173	0.342
Demand	1.97	2	0.148	0.062	< 0.0005	0.185	0.392
Skill	1.18	2	0.315	0.038	< 0.0005	0.146	0.248
Fit	1.84	2	0.167	0.058	< 0.0005	0.180	0.370

4.4 Affect

Participants completed the International Positive and Negative-affect Schedule Short Form (I-PANAS-SF) (see Section 3.3.2). The schedule consists of two five-item subscales: Positive-affect and Negative-affect. Each item is a scale of 1-5. Final scores are the sum of its item scores. See Table 4.8 p. 53 for descriptive statistics.

4.4.1 Main effects

Positive-affect data meets, or were corrected for, the theoretical assumptions for repeated measures ANOVA. Thus, we hypothesize that H_1: Policy has an effect on positive-affect, or H_0: policy has no effect on positive-affect. The data meets all of ANOVA's assumptions except for sphericity, where we used the Greenhouse-Geisser correction for degrees of freedom. There are no effects present here $F(1.549, 46.463) = 0.635$, $p = 0.496$, $\alpha = 0.05$ thus the data fails to reject the null hypothesis. The sample's data suggests that, contrary to our pilot findings, the size of the effect is small or medium. See Table 4.9 for details.

[26] Non-directional hypothesis, so critical $\alpha = 0.25$

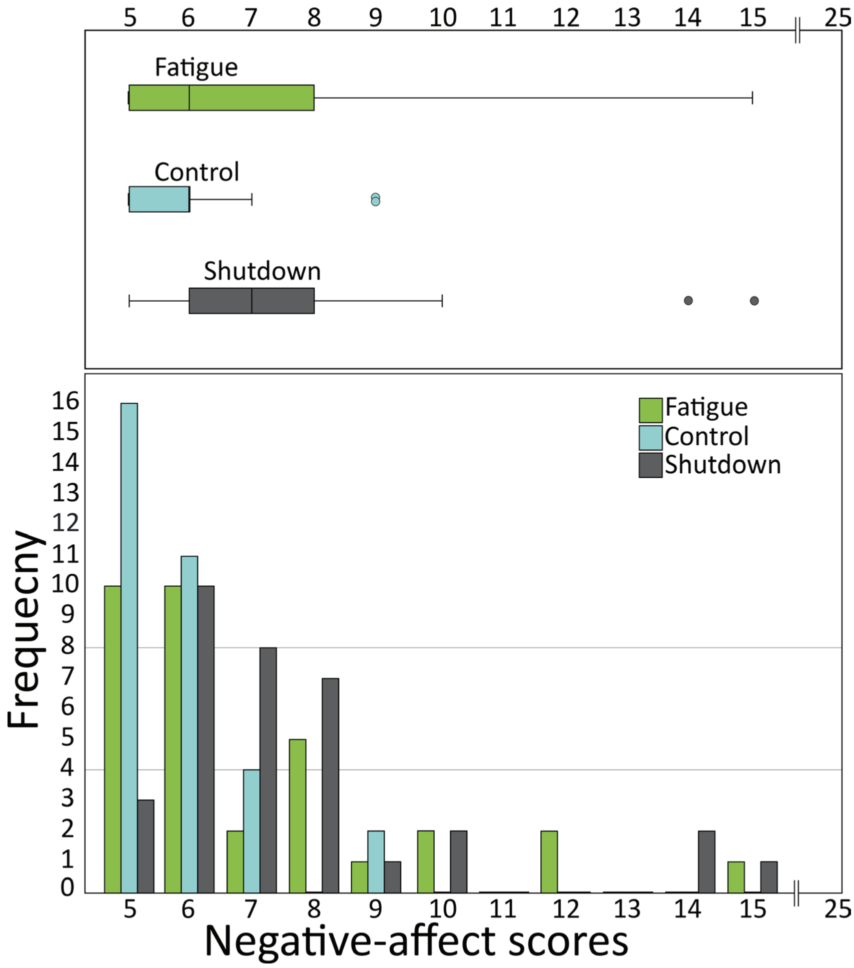

Figure 4.4 Frequency distribution of negative-affect scores.Scores on scale of 5-25; less is better. Each condition has 33 responses. The scores are clustered at the low-end of the scale with 27 of the control (blue) responses, and 20 of the fatigue (green) responses scoring 5 or 6.

Table 4.8 Descriptive stats for affect. Positive-affect scores (more is better) and Negative-affect (less is better) scores are on a scale of 5-25. The midpoint of the scale is 3. Totals based on all 99 data points, not condition values presented in table. Cronbach's α = 0.7 and 0.854 for PA and NA respectively based on our data.

						95% CI for Mean		
Factor	Condition	Number	Min	Max	Mean	Lower	Upper	SD
PA	Fatigue	33	8	24	14.85	13.34	16.36	4.27
	Control	33	8	24	15.61	14.13	17.08	4.15
	Shutdown	33	9	24	15.30	13.79	16.81	4.26
	Total	99	8	24	15.25	14.42	16.09	4.20
NA	Fatigue	33	5	15	7.04	6.16	7.90	2.44
	Control	33	5	9	5.82	5.44	6.20	1.07
	Shutdown	33	5	15	7.42	6.64	8.21	2.22
	Total	99	5	15	6.76	6.34	7.18	1.96

Table 4.9 Inferential statistics for PA (repeated Measures ANOVA).

				95% CI for η_p^2		
F	df	Sig. (p)	η_p^2	lower	upper	power
0.635	1.549	0.496	0.021	< 0.0005	0.137	0.138

The data for negative-affect is not normally distributed (Details in Table 4.10, Frequency distribution in Figure 4.4 on p. 53) therefore we use the Friedman test to identify effects, and absolute differences and confidence intervals to describe effect size. Our restated hypothesis that conforms to the outcomes of a Friedman's test: H_1: Sample distributions of Negative-affect are *different* for each policy, or H_0: Sample distributions of Negative-affect are *the same* for each policy.

Table 4.10 Tests of Normality for Negative-affect (Shapiro-Wilk test).

Condition	Statistic	df	Sig. (p)
Fatigue	0.791	33	< 0.0005
Control	0.735	33	< 0.0005
Shutdown	0.769	33	< 0.0005

The data shows a statistically significant difference in negative-affect, depending on the condition experienced, χ^2 (2) = 21.942, $p < 0.0005$, $\alpha = 0.05$. We reject the null hypothesis, and can say that the sample's distributions are indeed different, and that the population is very likely to exhibit similar behaviours.

Table 4.8 (p. 53) and Figure 4.4 (p. 53) presents the magnitude and range of scores for each condition. The scores recorded in the control condition are lower and consistent with a 75th percentile of 6, as opposed to 8 in both other conditions. The fatigue and shutdown conditions show a wider range of responses. Despite producing the same 75th percentile, the shutdown condition has smaller interquartile range, including a higher median than the other two conditions. This indicates that more participants reported higher levels of negative-affect for the shutdown condition.

Wilcoxon post-hoc test confirms that shutdown scores are significantly different from the control scores, Z = -3.693, $p < 0.0005$, $\alpha = 0.017$[27], showing a median increase of 1.5 points on the scale, 95% CI [1, 2]. It is likely that the population would experience an increase in negative-affect of similar magnitude. A significant difference exists between the fatigue scores and control scores, Z = -3.376, $p = 0.001$ $\alpha = 0.017$, but this time showing a median increase of 1 point on the scale, 95% CI [0.5, 1.5]. No significant difference exists between the fatigue and shutdown conditions.

Individual emotions

Negative-affect is a multi-factor construct and the I-PANAS-SF measures a single emotion for each item (see Table 2.2 p. 24 and Section 3.3.2). It is unclear from the total scores if a single item accounts for the observed differences. We performed a Friedman's tests on each item (see Appendix I.4 for details of all items tested), and found a significant difference for only the item measuring the *upset* dimension, χ^2 (2) = 28.659, $p < 0.0005$, $\alpha = 0.05$ (see Table 4.11).

---

[27] Whereas Friedman's test is robust, and accounts for this possibly, the Wilcoxon test does not and requires a more stringent significance level. We therefore apply the Bonferroni correction (α) to obtain critical $\alpha = 0.017$, using the following equation: $(\frac{original\ \alpha}{repetitions} \rightarrow \frac{0.05}{3} \therefore \alpha = 0.017)$

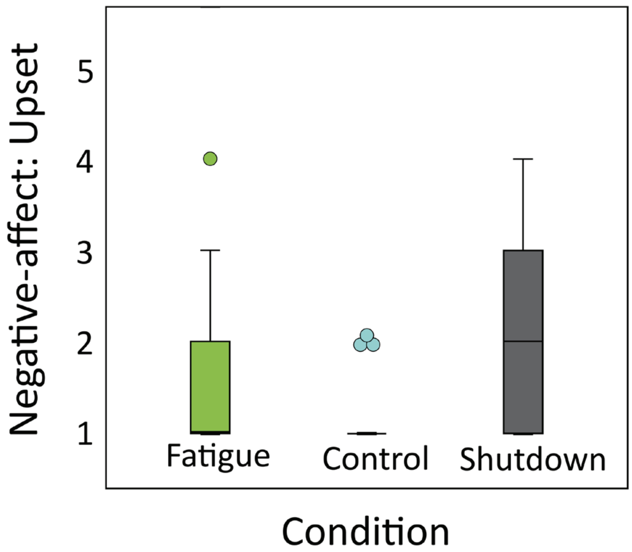

Figure 4.5 Sample distributions of Upset scores for each policy.Less is better. Circles are outliers. The control condition's spread is significantly thinner from those of the other conditions: 30 of the 33 instances recorded the lowest possible score for this factor.

Table 4.11 Descriptive and inferential statistics (Friedman's) for NA: upset only.

			IQR				Mean			
Condition	N	Min	25th	Med	75th	Max	Rank	χ^2	df	Sig. (p)
Fatigue	33	1	1	1	2	4	2.02			
Control	33	1	1	1	1	2	1.44	28.659	2	< 0.0005
Shutdown	33	1	1	2	3	4	2.55			

Wilcoxon tests of the *upset* dimension found a significant result between shutdown and control conditions, $Z = 4.185$, $p < 0.0005$, $\alpha = 0.017$, where the median shutdown scores are 1 point higher than median control scores, 95% CI [0.5, 1.5]. We also found a significant difference between fatigue and control conditions, $Z = 3.474$, $p < 0.0005$, $\alpha = 0.017$, where the median shutdown scores are 0.5 points higher than median control scores, 95% CI [0, 0.5] see Figure 4.5 for a distribution of responses. Although we observe that more players are upset after the shutdown sessions than they are after fatigue sessions, the difference is not statistically significant.

4.4.2 Carry-over, sampling and interaction effects

To test for carry-over effects, we group all data points that share the same ordinal value and compare them. In other words, we place the data from all first sessions and compare them with all data from second sessions, and third sessions. We found no significant effects between sessions with same ordinal value, $\chi^2(2) = 1.519$, $p = 0.468$, $\alpha = 0.05$ and infer that the carry-over effects are inconsequential to our data.

Table 4.12 Categories and descriptions of NA groups used for carry-over and sampling effect tests.

Effect type	Factor	Level 1	Level 2	Level 3
Carry-over	Ordinal value	First sessions	Second sessions	Third sessions
Sampling	Participant group	Sequence SCF	Sequence FSC	Sequence CFS

To test for sampling effects, we group all data points that share a particular sequence and compare those groups. In other words, we place all data that experienced the shutdown condition first in a group, and then compare it with the group containing all data of those that experienced the fatigue condition first, and the group containing those that experienced the control condition first (Table 4.12). Here we found no significant result between the participant groups, χ^2 (2) = 3.156, p = 0.206, α = 0.05. Thus, we failed to identify significant carry-over and sampling effects in our data for negative-affect.

4.5 Intentions to return

We asked players to indicate how soon they intended to return to the game. Participants indicated on a single item Likert-type scale with six options ranging from *never* to *immediately* (with values of one and six respectively). This information helps us understand how long the participants feel they will spend outside of the game itself.

4.5.1 Main Effects

The data is ordinal because the answer options indicate an ordered time value. However, the distances between responses are not uniform, so mean value testing may not represent the collective data. For example, a mean value of 3.5 may not express "between a 3 and 4" because that score is not possible on the scale. However, a median score of 3.5 does make sense: half the sample scored a three or less, while the other half scored a four or more. Thus, we use the Friedman test to analyse our data and measure the spread of responses across the scale, rather than mean-value testing. The Friedman test is better suited to categorical data. Our restated hypothesis that conforms to the outcomes of a Friedman's test: H_1: Sample distributions of intentions to return are *different* for each policy, or H_0: Sample distributions of intentions to return are *the same* for each policy.

We can see from Table 4.13 (p. 57) and Figure 4.6 (p. 57) that each condition's IQR the same width, but occupies a different position in the response space. The fatigue condition reports the lowest scores, followed by the control and shutdown conditions respectively. Furthermore, the IQR for the fatigue condition shifts down from the control by one point, whereas shutdown IQR shifts up by one point (see Table 4.5 p. 50).

Participants 32 and 33 are responsible for the two *never* entries. They both selected this during their first (and fatigue) session. These same sessions produced the two extremely long sessions (see Section 4.2). We cannot omit their other data based on one extreme value. Their subsequent responses fall within normal ranges for the respective conditions.

We found a significant difference between conditions, χ^2 (2) = 12.606, $p < 0.002$, α = 0.05. We reject the null hypothesis, and can say that the three condition's distributions are not all the same.

Table 4.13 Descriptive and inferential statistics (Friedman's) for Intention to return. Showing the range, relevant percentiles, and Interquartile Range (IQR). 1 = "Never", 2 = "Maybe never", 3 = "Sometime in the future", 4 = "Later this week", 5 = "Sometime today", 6 = "Immediately".

			IQR				Mean			
Condition	N	Min	25th	Med	75th	Max	Rank	χ^2	df	Sig. (p)
Fatigue	33	1	2	3	4	6	1.67			
Control	33	2	3	4	5	6	1.89	12.606	2	0.002
Shutdown	33	2	4	5	6	6	2.44			

Wilcoxon test confirms that the observed difference between shutdown and control scores is statistically significant, Z = -2.747, p = 0.006, α = 0.017, where shutdown responses are a median of 1 point higher on the scale, 95% CI [0.5, 1.5]. In other words, at least half of all responses will be one point higher for the shutdown condition than for the control.

Furthermore, statistically significant difference exists between the shutdown scores and fatigue scores, Z = 2.694, p = 0.007 α = 0.017, where shutdown responses are a median of 1.5 points higher on the scale, 95% CI [0.5, 2]. This indicates that half of all shutdown responses will probably be at least one point higher than fatigue responses. No significant difference exists between the fatigue and control conditions.

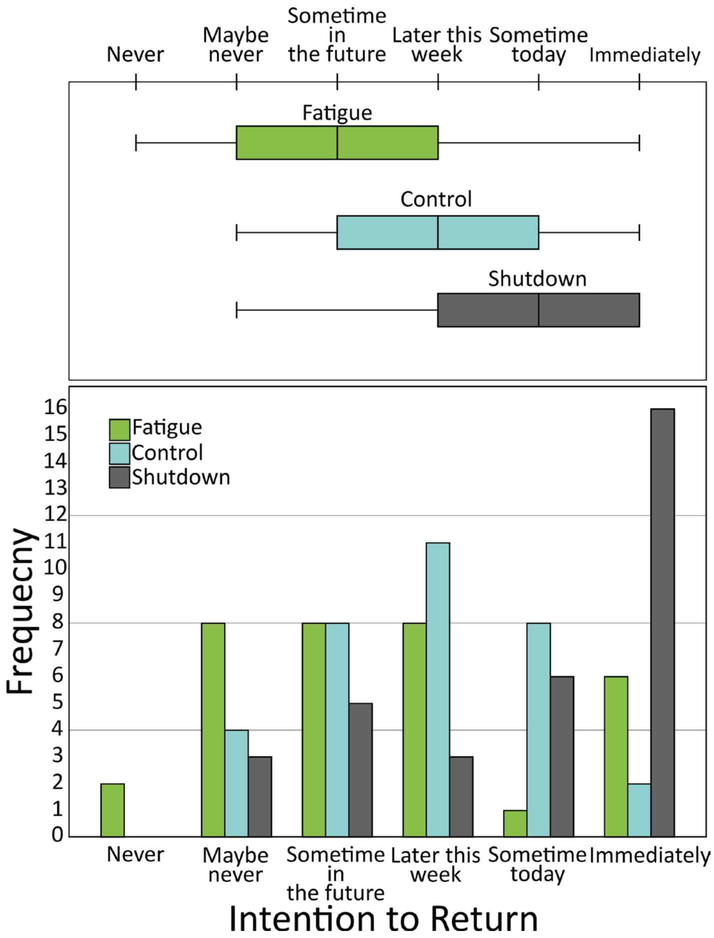

Figure 4.6 Sample distributions of intention to return to game.Frequency plot shows values and Boxplots elucidate percentile distributions. The fatigue condition shows a 75th percentile at "Later this week" whereas this option represents the 25th percentile for the shutdown condition, and the median for control. Sixty-six percent of shutdown sessions saw participants wanting to play again that same day, and nearly half wanting to play immediately.

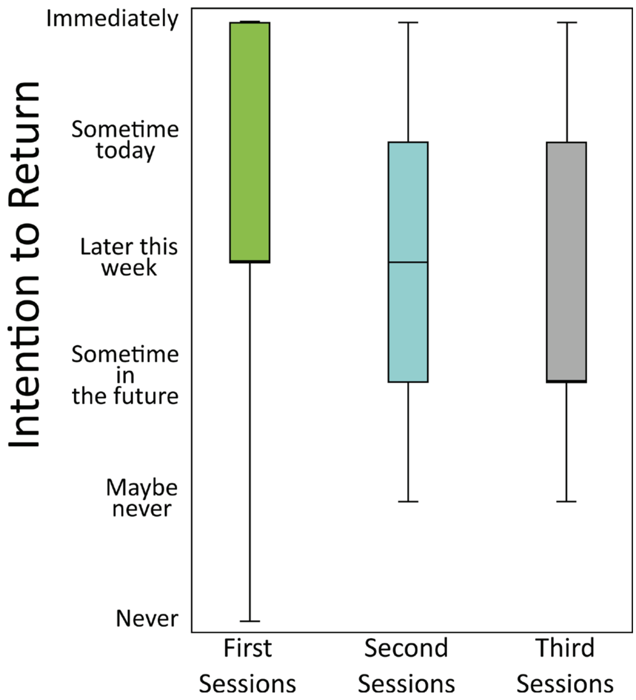

Figure 4.7 The spread of responses grouped by ordinal value.Players reported high scores more often after their first session, median = 4 IQR [4, 6] than their third Median = 3 IQR [3, 5].

4.5.2 Carry-over and sampling effects

To test for carry-over effects, we place the data from all first sessions and compare them with all data from second sessions, and third sessions. We then ran a Friedman's test on each category and found a significant effect for ordinal-value, χ^2 (2) = 6.165, p = 0.046, α = 0.05. Post hoc tests find the difference between first and third session only, Z = -2.457, p = 0.014, α = 0.017. Here the first session showed an increased median of one point on the scale, 95% CI [0, 1.5] (see Table 4.14 and Figure 4.7 on p. 58).

Table 4.14 Wilcoxon tests for ordinal values of intentions to return. Each row shows the pair of ordinal-values being compared.

Comparison	Median difference	Sig. (p)	95% CI for difference	
			Lower	Upper
First – Second	-0.5	0.356	-1	-0.5
First – Third	-1	0.014	-1.5	0
Second – Third	-0.5	0.178	-1.5	0

To test for sampling effects, we group all data points that share a sequence and compare those groups. In other words, we place all data that experienced the shutdown condition first in a group, and then compare it with the group containing all data of those that experienced the fatigue condition first, and the group containing those that experienced the control condition first. We then ran a Friedman's test on each category and found no significant result between participant groups χ^2 (2) = 0.58, p = 0.971, α = 0.05. Thus, we failed to identify significant effects due to sequence allocation, which suggests that sampling effects are inconsequential for our data.

4.6 Player's reason for hiatus

We asked players to detail why they chose to end each game session. For the shutdown condition, the reason was always the same: the game terminated automatically. We recorded 33 responses in the control condition. One participant submitted "violent" as their response for the fatigue session, while another submitted no response for the same condition; therefore, we have only 31 data points for the fatigue condition. Players related a variety of contributing factors, which we subject to thematic analysis. See Appendix J for raw responses.

To answer research question 4

> *How does the design policy change a player's reason for choosing to end their session?*

we used a grounded approach to thematic analysis and identified common and recurring themes in the set of responses. We began by isolating responses from any identifying information. We then pooled and randomized the fatigue and control responses. The primary experimenter began with a blind read of each response sample wide in a random order to minimize experimenter biases. The reader looked for punctuation, capital letters, large whitespace, and conjunctions to identify complete thoughts, and thereby reduce each participant's response to a set of utterances: a single word or phrase that captured its essence as a reason for hiatus. Recurring patterns emerged from these utterances, which we identified and categorized into themes that specifically relate to the participant's reasons for hiatus.

We reduced themes to three encompassing categories. An undesirable experience or *discomfort* captures all responses wherein participants became aware of some in-game factor that made continued play laborious, uncomfortable or undesirable. In the fatigue condition, this includes the loot-fatigue mechanics. A *satisfied* hiatus, captures the responses wherein participants felt their desire to play was satiated. The term *other priority* captures the responses wherein participants had another commitment that required their attention, and they could play no longer without suffering some undesirable outcome. We then re-read each response at the condition level and placed it in one of these categories. See Figure 4.8 for the frequency distribution of categories.

In both conditions, the most common reason for ending a session was due to the onset of an *undesirable experience* in game. The second most common reason is some other, time critical priority that compels players to stop. This is not to say that discomfort, as a reason for hiatus, means players had a negative experience, just that the cost of continuing outweighed the benefits of playing. The least common reason why players stopped was that they felt satisfied. In the control condition, *other priority* occurs an additional four times while there were four fewer instances of undesirable experiences. Furthermore, nearly half of the responses in the undesirable experience/discomfort responses explicitly cited the fatigue system as a reason for hiatus.

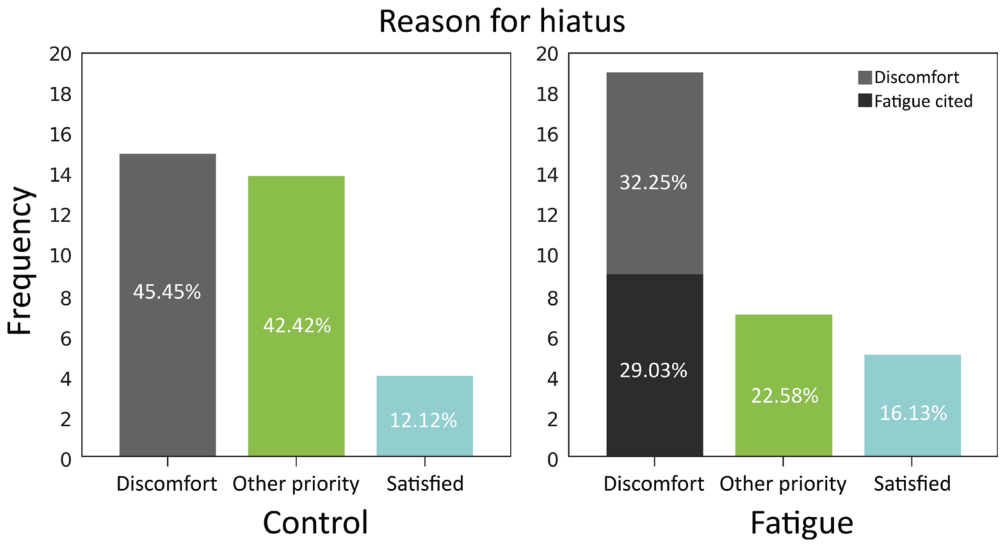

Figure 4.8 Frequency distribution of participants' reasons for stopping in the control and fatigue conditions. Undesirable experience/discomfort describes players stopping because the game began providing them with negative experiences. Satisfied describes players stopping because they felt satisfied. Other priority describes stopping for reasons unrelated to the game. The green discomfort results indicate players who explicitly cited the fatigue condition as part of their reason for hiatus.

4.7 Dimension correlations

Finally, we computed Pearson’s correlations for all our relevant dimensions in Table 4.15.

Table 4.15 Pearson’s Correlations between all factors. Highlighted cells are significant correlations. Red highlights indicate established correlations due to their being a part of same scale. Green highlights represent new correlations.

Abs = absorption, Perf = performance, Imp = importance, Dem = Demand, Fit = fitness, Int = intention to return.

		Flow	Abs	Perf	Imp	Dem	Skill	Fit	PA	NA	Time	Int
Flow	Corr.	1.00	.847	.913	.439	.009	.385	.155	.411	-.010	-.299	.212
	Sig.	1.00	.000	.000	.000	.926	.000	.126	.000	.918	.003	.035
Abs	Corr.		1.00	.556	.461	.167	.227	.324	.369	.064	-.264	.287
	Sig.		1.00	.000	.000	.099	.024	.001	.000	.526	.008	.004
Perf	Corr.			1.00	.334	-.113	.425	-.007	.360	-.066	-.263	.111
	Sig.			1.00	.001	.264	.000	.945	.000	.518	.009	.275
Imp	Corr.				1.00	.171	.159	.387	.388	.304	-.147	.259
	Sig.				1.00	.091	.117	.000	.000	.002	.148	.010
Dem	Corr.					1.00	-.061	.488	.108	.223	-.123	.090
	Sig.					1.00	.547	.000	.285	.027	.226	.375
Skill	Corr.						1.00	-.085	.107	-.093	-.253	.169
	Sig.						1.00	.403	.292	.358	.011	.094
Fit	Corr.							1.00	.163	.290	.024	-.067
	Sig.							1.00	.107	.004	.815	.513
PA	Corr.								1.00	.213	-.106	.362
	Sig.								1.00	.034	.295	.000
NA	Corr.									1.00	-.010	.011
	Sig.									1.00	.920	.910
Time	Corr.										1.00	-.440
	Sig.										1.00	.000

All significant correlations discovered are weak to moderate in strength. We discuss only those with a magnitude of 0.4 or greater. We found a moderate positive correlation Flow and positive-affect suggesting that as flow increases, so does the experience of positive-affect. This result is unsurprising given the rich discourse surrounding flow as an optimal experience contributing to positive well-being.

We found a moderate negative correlation between playtime and intent to return to game. This suggests that the longer players spend game, the less they desire to return. This corroborates our discussion of intent in Section 4.5

5 DISCUSSION

5.1 Introduction

In the previous chapter, we presented the data we collected, and the results of all statistical analyses. In this chapter, we discuss what those results mean for our research.

5.2 Participants

Our participants played our experimental system, built in Torchlight II, for an average *and* median of 4.85 hours over three sessions, placing them in the 30th percentile of all players worldwide (Steamspy, 2017); the global median for total time spent playing Torchlight II is 11.25 hours (Steamspy, 2017). Although they did not complete the game in any way, our participants played a significant portion. This time spent in game and the high flow scores (see Section 4.3.1) suggest that our participants were intrinsically motivated to play the game. Our participants further expressed high positive affect and generally low negative affect scores in keeping with the notion that games have positive effects on emotions. Thus, we are confident that participants provided data that probably represents population data of players who have invested in a game.

5.3 Session lengths

The behaviour policies are represented by the shutdown condition that limits session length to one hour. The design policy is represented by the fatigue condition in which loot [28] becomes increasingly rare after one hour of play. We hypothesized that

[28] Gold, Magic items, and character's experience.

behaviour and design policies influence session length. Our data strongly supports this hypothesis

5.3.1 Shutdown condition

We found the shutdown condition reduces the time that players spend in game when compared to our other conditions. Our data shows that 32 of our 33 participants played for more than an hour in the other conditions[29]. This behaviour suggests that one hour is less than the average desired session length for our chosen genre (action RPG), game (Torchlight II), in context (Single-player, offline, non-leisure laboratory setting). They are unable to play for longer under the condition of the behaviour policy. Seventy-five percent of control session were less than two hours; the effects of a behaviour policy on session length will change depending on the period given to players. For example, if we allowed a two-hour period then most players would not have experienced the shutdown event even if their contextual factors such as free-time allowed them too. That being said, our methods required that players be made aware of their time limit, framing it as a scarce resource.

The Taiwanese policy suggests session lengths of no more than half an hour at a time (Locker, 2015; Seok Hwai, 2015); it is unclear if there is a daily limit. Our data suggest that the smallest period players spend in game is one hour; half that period would certainly affect player behaviours. Further research is required to observe what these might be. For all periods, knowledge of the time limit may affect the goals that players set for themselves, especially if they know they cannot accomplish a goal within the allotted period.

5.3.2 Fatigue condition

Our data show that the average participant is willing to play for two hours with half of that time subject to fatigue mechanics. This result suggests that the modified mechanics do not frustrate players enough to inspire an early exit from the game and that something other than loot mechanics motivate continued play. Storyline progression commonly motivates play in the early stages of a game. In our study, it is possible that the desire to experience the story motivates continued play more than character progression. For example, participants *R* and *V* both expressed the desire to complete one more quest before ending their session, despite:

R: I'll just do one more quest.

and

After an hour and thirty-nine minutes of play:

V: "I just want to play forever! I keep dying....must I stop?"

[29] The only non-shutdown participant that played for exactly one hour indicated their intention and set an alarm

Experimenter: "It's up to you entirely"

V: "okay, I just have to finish this quest"

And then *V* continued playing for an additional thirty-seven minutes adopting a slow, more conservative, ranged battle strategy to avoid damage as much as possible to complete said quest without dying as much. One reason *V* struggled here was due to a lack of potions; a potentially common challenge created by the fatigue mechanics. This phenomenon was identified by participant *R* around two hours into their fatigue session (one hour after the fatigue mechanics took effect):

R: Potions are so expensive now!

Potions in Torchlight II are low-cost consumable items that sustain a character during battles by restoring health, among other things. Players may purchase and consume hundreds of potions in one session. However, as gold becomes rare, it becomes more difficult to keep buying more. *V* had already spent thirty-nine minutes under the effects of the fatigue condition, gaining no experience, finding inconsequential amounts of gold, and receiving few items to sell for potion money.

Another reason *V* struggled was because their character had not levelled up, and grown stronger relative to the enemies they were fighting. Not only does this contribute to increased risk of death, as participant *J* indicated:

J: ya dude, I'm still level 9. I don't think I can compete. I'm going to get wrecked.

It also increases the time required to play the game. Using a simple example, a level 10 character must attack a level 10 enemy 5 times to kill it, but a level 8 character must attack a level 10 enemy 7 times to kill it. The time needed to perform extra attacks adds up when the game and genre ask players to kill hundreds of enemies per session.

Players may not care about the mechanic impediment if their goal is to reach the next story event, or complete quests. In this case, the system may be unwittingly facilitating Chen's flow model (Chen, 2007; see also Section 2.4.1), wherein the game offers players a choice of challenges: continue exploring new areas despite the increased difficulties. Given these fresh gameplay options, they decide for themselves to continue playing without character progression for as long as gameplay remains challenging and stimulating.

In fact, any goal that players set for themselves may temporarily replace character progression as the player's priority without their knowing it, as was the case with participant Y:

Y: The most frustrating things was when the XP stopped entirely. I only realized half an hour after [the fatigue began] that it had stopped. I was so close to levelling up that I kept saying 'one more area, one more area' and I never got there! It was the most frustrating.

Here, Y appears to have set them self the short-term goal of completing areas of the game; a goal they chose to complete while levelling up. This distracted them from character progression and motivated continued play, thereby increasing session length.

It might seem that participants do not care for loot at this stage of the game, but there are those that do. For example, participant *F* swore loudly upon finding out about the fatigue mechanics. Participant *H* expressed their dilemma when tasked with a difficult maze:

H Enters maze challenge

H: [exclaims] Whyyyy?! Ugggh. I don't want to.

Experimenter: You can leave, right?

H: Ja, but…the treasure… This is my least favourite challenge.

Some players may try to offset the fatigue mechanics, using their skill and knowledge of game mechanics to keep the loot rates up. For example, the strategy relayed by participant *F* in the fatigue condition:

F: Every time I found items [wearable loot] that would increase magic drop [rates], gold drop [rates] or XP [gain rates] I immediately put them on. Screw health [laughs]

Participant F attempted to offset the effects of the fatigue condition. They perceived that the condition reduced their freedom to play as they wished, and they reacted against it. They placed more value on the opportunities to keep finding loot than the opportunity to keep playing. In doing so, they set themselves a personal goal for the session: Find loot that increases the chance of finding more loot. In other games with different implementations fatigue this strategy may work, however our implementation does not allow for it. It would only be through trial and error that these players would see the folly of this strategy, and one session was not enough for them to notice this.

Players not as resourceful as participant F may also be subject to a planning fallacy, whereby they are unable to foresee that the fatigue mechanics will impede their progress at later stage of the game. The following sessions may have their progress restricted for some time because of resource strain placed on the previous session: players may not have a store of potions, items, or gold required to progress in the game, and will need to spend time earning gold to afford the potions required.

Fatigue compared to control

We found that players spent 24 minutes more playing the fatigue condition (M = 127) than they spent playing the control condition (M = 103); we expect the true difference to between 9 and 38 minutes based on the confidence intervals (see Table 4.2 p. 47). This result suggests that exposure to design policies increases session length enough to be *noticeably longer* than control sessions to anyone paying attention (Cohen, 1988). We speculate from this observation, and the comments made by participants, that by restricting character progression the game demands more attention and skill from players sooner than intended by the game's designers. For example, less gold-income means players need to manage the game's internal economy better lest they run out of money for important items such as health and mana potions. Our participants played Torchlight II during the early game where unaffected characters make significant progress several times per area. Fatigued characters would make no effective progress per area; their abilities no longer scale with the abilities of their enemies. Players familiar with the genre, such as our participants, can overcome the additional demands, albeit for a limited period. They compensate for the relative weaknesses generated by the fatigue system with specialized knowledge of the game including enemy weaknesses, and their mastery of game mechanics such as player positioning and their pet's behaviour. Perhaps most importantly, participants use extra time to progress because players may need to attack a monster for longer periods to defeat it, or return to town more often. This effect is not limited to goals of character progression. Practically all goals that players set for themselves in Torchlight II become more difficult to accomplish because the fatigue condition increases the effort required in some way, be it cognitive attention or time required. It is certainly possible that by inflating the game's demands, the design policy manipulates the player's perceived balance of demand and ability (see Section 2.4.1 and Figure 2.1 on p. 20), thereby influencing flow experiences and the perception of time passing. We explore this idea further in Section 5.4.

We are confident that a gamer playing under the effects of the fatigue condition is highly likely to play for longer than if they played under our control condition. This along with the differences observed in mean session length suggest that the condition influences the session length.

5.4 Flow

Our results suggest that our system facilitates flow experiences. Using the Flow Short Scale (FSS) we measured a mean of $M = 5.29$ (0.94) for *flow*, which describes *fluency of performance* and *absorption*, and $M = 3.53$ (1.59) for *importance*, which describes how much value our participants placed on the outcome of the activity. Whereas the metric does not provide threshold scores describing when the flow experience becomes strong or deep our results are high relative to other data sets. Rheinberg and colleagues (2005, as cited in Rheinberg, 2015) collected 4479 FSS results across a variety of activities recording a mean of $M = 4.97$ (1.18) for flow and $M = 2.26$ (1.54) for importance (see Table 5.1 p. 69 for select activities; For full list, see Rheinberg, 2015). They do not describe the three remaining FSS items: demand, skill, and the perceived balance between them. Thus, we cannot compare them to our results.

We hypothesised that policy has an effect on flow scores. We cannot reject the null hypothesis for any dimension of the FSS. All observed effects are non-significant, and are small to moderate in size. This result does not allow us to infer a zero effect for the population. In general, the statistics tell us very little about the data. The data suggests a moderate effect size ($\eta_p^2 = 0.09$), but also shows a wide confidence interval, thus we cannot suggest with any certainty if the true population effect size is small, moderate or large. Furthermore, the *observed* power is well below the expected 0.8, which is surprising especially considering the relatively low p-values[30]. That being said, our results do not appear to be confounded by carry-over or sampling effects

It is possible that the general population is too diverse to identify significant effects through our methods or with our sample because our pilot sample is a poor reflection of the population, leading to inaccurate sample-size estimations. Furthermore, we did control *game* and *genre*, but did not control for, or correctly identify, all aspects of the *player* variables, including player types and motivations. Finally, we were unable to control for temporary player contexts, such as their state of mind, and other factors affecting their ability to experience flow during a session. Further experiments would require more participants, or strict categorization according to a player typology, to confirm an effect and provide better understanding of the effect size.

We expected the scores for the FSS's *importance* dimension to rise with each consecutive session as players invested more time into their character – as much as seven hours. We found no evidence for this; importance scores remained similar for each subsequent session.

We expected to see a difference in demand scores, specifically an increase during the fatigue condition because of the impeded character progression. Overall, demand is low, skill is high, and the perceived fit is middling. The evidence strongly suggests that players did not feel that the system demands more skilled play or cognitive effort to overcome the fatigue mechanics. However, the weak negative correlation between session length and perceived skill suggest that those who felt less skilled at the game also played for longer periods. Considering our findings for session length (see Section 5.3.2), this evidence suggests that the fatigue mechanics degrade the flow experience. If anything, the evidence suggests that demands of the game change, but maintain their intensity; players are aware that rules have changed, and that their character will die more often (see player comments in Section 5.3.2), but they continued to play regardless and required only one additional resource to reach their goals: time.

Whereas the conditions themselves failed to affect the FSS results, comparisons between our results and those of other activities proves interesting. Not all the activities listed in Table 5.1 p. 69 occupy same context commonly associated with video game play (sleeping, moving, and routine work). Thus, we selected *working with a PC, goal directed communication, watching TV, reading, sports, recreating (sic. means "doing" recreational activities),* and *gardening* and found means of M = 5.01 (1.14) and M = 2.10 (1.25) for Flow and Importance respectively. Our data

---

[30] Observed power is a function of p values. As p-values tend towards zero, observed power tends towards 1 (Hoenig & Heisey, 2001).

presented a total mean of $M = 5.29$ (0.94) for flow and $M = 3.53$ (1.59) for importance. A one-sample t-test found a significant effect; our results are higher for both components (see Appendix I.4).

Table 5.1 Descriptive statistics gathered for the Flow Short Scale in other activities. Our results are presented in the top line, the Pac-Man result are Engeser and Rheinberg (2008), All other activities are from Rheinberg and colleagues (2005, as cited in Rheinberg, 2015).

	Flow		Importance	
Dataset	Mean	SD	Mean	SD
Our results	5.29	0.94	3.53	1.59
All activities (reported)	4.97	1.18	2.26	1.54
Select activities[31] (calculated)	5.01	1.14	2.10	1.25
Creating/inventing something	5.30	1.11	3.53	1.61
special activities on the job	5.59	0.90	3.76	1.56
doing office work	5.03	0.78	3.47	1.89
Studying for a statistics test	4.60	1.16	3.45	1.44
Pac-man Medium time 1	4.68	1.18	1.65	0.86
Pac-man Medium time 2	5.21	1.03	1.43	0.83
Pac-man very difficult	3.08	0.69	-	-
Pac-man very easy	3.83	0.92	-	-

Engeser and Rheinberg's 2008 study saw participants playing four differently difficult games of Pac-man: one very easy game, two medium difficulty games and one very difficult (see Table 5.1). They completed the FSS after every game. The report shows fair-to-middling flow and importance, all of which are significantly lower than our means with one exception – the second *medium* difficulty instance (see Table 5.1).

These comparisons suggest that players experienced high levels of flow regardless of the condition experienced. Our attempts to manipulate the experience of flow through the fatigue system are either unsuccessful, or confounded by other variables such as renewed interest in the novel and different challenges raised by stifling loot.

We expect that the population most affected by policy value gaming as an activity, and that they are highly motivated to play games. Our participants recorded importance scores (the items assessing how consequential a person perceives the activity and its outcome to be) are high relative to the majority of other activities reported on by Rheinberg and colleagues (2015; see Table 5.1). This, together with the flow results imply that games may offer autotelic experiences, or that flow increases the importance that players grant the activity. It is difficult to know if these importance scores facilitated flow, or resulted from flow.

Although we are unable to conclude a zero effect for our specific set of conditions, our results scores suggest that the population's experience of flow will remain high despite interventions such as these, when considered in context of the pre-existing reference scores.

[31] Activities include Working with a PC, goal directed communication, watching TV, reading, engaging in sports activities, relaxing/recreating and gardening.

5.5 Affect

5.5.1 Positive-affect

We found no evidence to support our hypothesis that the policies influence positive-affect. We observe a small effect size suggesting a practically null-effect in our sample however, the confidence interval is wide. The true effect may be small, moderate or large. We are thus unable to prove a practically null effect between our conditions for the population.

5.5.2 Negative-affect

Statistical analyses of our data support the hypothesis that policy may increase negative-affect in players specifically by making players more upset. The data shows the largest effect in the shutdown condition where most participants experienced more negative-affect than not. We cannot prove that carry-over and sampling effects are absent from our data through hypothesis testing, or other means. However, we can infer that they probably do not influence the results in a meaningful way. Thus, we are certain that the shutdown and fatigue conditions increase negative-affect scores relative to control conditions - especially the feeling of being upset.

Behaviour policy

The shutdown condition removes player agency and they are unable to make decisions about their session length beyond a certain period. As the game session shuts down, players lose personal control when the freedom to keep playing becomes unattainable. Brehm and Brehm's *psychological reactance* theory (1981) describe how when personal freedoms become limited, the need to retain those freedoms makes people desire them more despite their utility. It is probable that players report negative affective-responses when they are unable to use their time effectively and complete personal goals. Participant *B* expressed this idea after their control session, which followed their shutdown session:

> *B: This time I actually managed to complete a quest before finishing.*

This suggests that participant *B* was conscious of their being unable to complete a quest the previous session. Players may perceive a loss of personal control over their behaviour and become upset when they find no way to regain control. Finally, the strict limit may reduce digital games' efficacy to offset players' negative mood.

Design policy

Negative-affect increases are small in the fatigue condition. The policy did produce the widest distribution of scores within its range; it has the largest IQR. This indicates that the fatigued responses are often higher than the scores recorded after the control condition, but the effect size suggests that the difference may not be practically meaningful. The fatigue IQR is the union of the shutdown and control IQRs. Thus, we are not able to identify the policy based on score alone because a high score may result from a behaviour or design policy, while a low score may result from the design policy or no policy (control). It may be that some participants experienced no increase in negative-affect because they were able to accomplish their personal or game-set goals despite the fatigue mechanics. Some participants may not have possessed the ability or desire to forge through the impediment and felt disappointed when they did not accomplish their goals, thus resulting in increased negative-affect. Either way, the fatigue system produces inconsistent change.

5.5.3 Result in context

We must also consider that our data are relative to our control condition only. The scores remain significantly lower than those reported as traits experiences by Thompson (2007), and are similar to the means from other studies. By these accounts playing games decreases the experience of negative-affect, even when authorities police hiatus behaviour.

Table 5.2 Comparison of our affect results with Thompson's (2007) means, and meta-mean. Meta-mean averaged from a meta-analysis of research using the instrument as a state measure (Falconer et al., 2014; Fredenburg & Silverman, 2014; Karim et al., 2011; Kuesten et al., 2014; Sanchez et al., 2014; Vermeulen et al., 2016), and the scores from our research.

			95% CI for Mean		
Factor	Study	Mean	Lower	Upper	SD
PA	Thompson	19.15	unknown		2.77
	Meta	14.62	unknown		2.88
	Our results	15.25	14.42	16.09	4.20
NA	Thompson	12.73	unknown		3.01
	Meta	7.06	unknown		1.56
	Our results	6.76	6.34	7.18	1.96

Table 5.2 displays descriptive stats for several other studies using the I-PANAS-SF. Thompson (2007) reported scores for communities across the world when they developed and validated that I-PANAS-SF. Our sample exhibited lower and significantly different affective response in both constructs and for each condition (see Appendix I.4). This result is at first concerning: the discrepancy immediately suggests that our versions of Torchlight II produce a flattened affect; a low emotional response. As mentioned in Section 3.3.2, Thompson (2007) collected scores using the questionnaire as a trait measure, meaning that their scores reflect how people generally feel in their daily lives, not after a specific event. It would be unwise to suggest that playing our system may reduce positive affect compared to normal scores We found research using the I-PANAS-SF as a state measure reports similar scores to our own, where the mean falls within our mean's 95% confidence interval (Falconer et al., 2014; Fredenburg & Silverman, 2014; Karim et al., 2011; Kuesten et al., 2014; Sanchez et al., 2014; Vermeulen et al., 2016). We are confident that our task utilizes the I-PANAS-SF in a normal way. See Appendix I.4 for details of all tests involved.

5.5.4 Conclusion

Games are used for emotion regulation, not necessarily to increase positive-affect, but to offset negative emotions and reduce the negative-affects (see Section 2.2). Both behaviour and design policies appeared to increase negative-affect by a small amount, but there is no apparent decrease in positive-affect. It may be that these policies do not increase negative-affect per se, but rather decrease the medium's efficacy in offsetting negative moods – players do not become more upset, but rather remain upset.

It would be naïve to suggest that session length contributes to how upset players were. One could argue that the shutdown sessions were too short to allow the benefits to manifest in each player; however, the fatigue sessions were long enough to incorporate an entire control condition's worth of mood management, and still it resulted in higher negative-affect. Rather, digital games offer players opportunities such as decision-making and strong personal agency that are hindered by the policies; the shutdown condition removes these entirely, whereas the fatigue conditions obstruct the game mechanics that provide such opportunities.

Whether the policies create negative moods, or impede player ability to offset negative mood, the result is that games lose one of the most vital aspects, and much like a self-fulfilling prophecy, they become that which policy makers believe them to be: a waste of time.

5.6 Intentions to return to the game

The results paint a neat picture of how policies may affect intentions to return. Our sample's responses conform to our expectations. Most players wish to continue playing immediately after the shutdown condition, whereas more players intended to take a longer hiatus after the fatigue condition than both others. Some participants expressed this, such as participant *L* after their fatigue session:

> *L: I don't feel as excited to play again as I did the last two times. Maybe I need to learn some new spells [character skills] or something. I felt like all I was doing was clicking and hitting everything until it died."*

The value from the control condition fits in between these two.

We found evidence to support our hypothesis that both policies probably affect intentions to returns in the population and that the shutdown condition is far more likely to do so than the fatigue condition. It is therefore possible that the shutdown condition affects players in a different manner to those created by the other two conditions.

Some responses appear anomalous. In both the control and fatigue conditions, some participants indicate that they wished to return to the game immediately after choosing to stop, which suggests that players ended their session despite wanting to continue. Participants were thoroughly enjoying themselves, but felt obliged to end their session due to hunger or other commitments (see Section 4.6 for more about the reason why players ended their session). Although these players chose hiatus, they did so reluctantly, which affected their intentions for future sessions.

From this and the significant result for the shutdown conditions, we can infer that players plan to return to the game sooner when forced to end their sessions for any reason. These results suggest that when players are not satisfied or satiated from their session they may plan to behave differently during their hiatus. For example, they may plan to play again sooner and displace another activity, or other less important commitment.

Significantly, more players indicated that they wished to return sooner after their first session than after their third session. Given the nature of our design, and the commitment required from participants, their knowledge of it being their final session with the game, may have affected their commitment to the game during the final session. This may have affected responses, but our counter-balanced design probably offset the effect enough to not confound our findings for policy. Nonetheless, this finding is worth discussing. A player's commitment to a current session affects their own forecast for their own behaviour, a forecast they no doubt use to plan their lives to some extent.

Finally, we acknowledge that this measure is not a psychometrically valid item measuring motivations, nor does it provide any indication of actual hiatus length; players may change their intentions over time and return to the game sooner than they intended after their previous session. Thus, we must be careful when drawing conclusion from this data.

5.7 Reasons for hiatus

We categorized most of responses (34 out of 62) as a form *discomfort*. These players became aware that continued play would perpetuate some form of discomfort, such as tiredness, hunger, or boredom that they are unwilling, or unable to ignore. In Section 2.4.2, we discussed how people in a flow state may not be aware of such discomforts because their attention is devoted to their current activity. If our data is representative of the population, then the fatigue condition appears to delay awareness of these discomforts by twenty-four minutes on average. Participants are at least willing to ignore the discomfort until they reach a suitable exit point such as completing a quest. This suggesting that most players will continue to play until the game no longer offers them the stimulating, positive, and desirable experience, and players become more aware of any discomforts. These players do not moderate their gaming behaviour per se, but rather choose to replace gaming with another occupation to better satisfy their current needs. This distinction is important. It relates to the notion that problematic gamers have little meaningful occupation in their lives besides playing digital games, or that other activities generate more discomfort for them than do games. In these situations, they will continue playing despite what most would consider discomfort, for example tiredness and boredom (Kowert & Quandt, 2016).

The responses that fall into the *other priority* category suggest similar intention. Players do not choose to stop playing, but are compelled to pursue another activity; often some impending commitment. Failure to honour those commitments may have caused great distress for the players, in the form of academic failure, or social conflict. In these instances, players stopped to avoid distress. As with discomfort, players who suffer from problematic videogame playing may not have many social commitments.

The instances where players stopped because they felt *satisfied* are the least common. Most of the responses suggested that the players had met a goal imposed by the game or themselves, and that continuing would require renewed commitment to the activity.

Having knowledge of each participant's reason for ending their session proved useful when considering intentions to return. It explained why players who seemingly chose to stop playing also desired to continue playing immediately (see Section 4.5.1). That insight alone gives cause to consider that the way in which games sessions end may contribute to different game seeking behaviours. Despite maintaining the positive aspects of UXG, the hiatus event may have created some level of aversion in players.

5.8 Summary of each policy and conclusion

Our experiment measured many aspects of behaviour and design policies. We found that the shutdown condition significantly reduces session length, and increases (fails to decrease) negative-affect – particularly the *upset* dimension. Finally, players want to play again sooner when they are forced to stop. The shutdown condition does not influence the positive aspects of UXG, namely flow and positive-affect.

We found that the fatigue condition significantly *increases* session length but does not change any aspects of UXG in practically meaningful ways. Here, players indicated that they intend to take more time between sessions than after the other conditions. Players were more likely to stop playing because of something related to the game that created discomfort.

6 Conclusion

This dissertation presents research about the effects that session length policies have on the gaming experience of regular players. We began by describing digital games in terms of their prevalence and their current popular moral debates before presenting empirical evidence surrounding various debates. Through this, we identified that subjective user-experiences (UXGs) are an important factor contributing to the effects of digital games. We described two common policies that authorities use to minimize the purported negative effects of gaming, primarily by reducing game session length. The behaviour policy (see Section 1.2.1) strictly limits session length, whereas the design policy (see Section 1.2.2) fatigues game mechanics ostensibly encouraging shorter sessions. The literature shows that positive and negative effects exist but only a small group of players are at risk; where authorities enforce session length policies, all players are affected. Thus, policy may degrade the intended subjective UXGs and possibly negate positive effects in the population. We built a system to incorporate these two policies and ran an experiment to compare the behaviour policy, the design policy, policy-free sessions, represented by our shutdown, fatigue and control conditions. We analysed our results to identify which policy best maintains the positive effects (see Section 2.2) of gaming, while minimizing the potential for negative effects (see Section 2.2.2).

6.1 Summary of achievements and results

Our research asked four questions of the session length policies. See Section 1.3 for a detailed discussion.

1. How does policy change a player's session length?
2. How does policy change a player's subjective user-experience?
3. How does policy change a player's intention to return to the game?
4. How does the design policy change a player's reason for choosing to end their current session?

Table 6.1 Summary of statistical methods used and results.

Factor	Significant result (dimensions)	Effect size[32] (absolute difference)	Conditions involved
Session length	Yes	Large (~23 minutes)	Fatigue > Control > Shutdown
Flow	No (0/6)	Inconsequential	None
Positive-affect	No (0/5)	Inconsequential	None
Negative-affect	Yes (1/5)	Small (1 out of 25)	Shutdown > Control
- Upset dimension	Yes	Small (1 out of 5)	Shutdown > Control
Intention to return	Yes	Small (1 out of 6)	Shutdown > Fatigue
Reason for hiatus	Yes	NA	Fatigue and Control

To answer these questions, we asked all participants to play all three versions of our ad-hoc experimental system using the critically successful and content rich action RPG, Torchlight II. Our system successfully allows players to experience what it must be like to play a game while subject to the three types of policy. Through a novel experiment comparing all three policies, we measured the length of each session, and the subjective user-experiences of each player that we operationalized in terms of flow and psychological affect. We utilized the Flow Short Scale (FSS; Rheinberg, Vollmeyer & Engeser, 2003; cf. Engeser & Rheinberg, 2008)) and the International Positive and Negative-affect Schedule Short Form (I-PANAS-SF; Thompson, 2007) for measurement. After each session, we asked players how soon they intended to play the game again. We asked players why they stopped playing when they did and performed thematic analysis on their answers.

Our experiment produced a large amount of data and offers new insight into how these common methods for reducing screen time affect player behaviours and experiences, and if they work as intended. See Table 6.1 for a high-level summary of our results. In general, we found that each policy has a large unique effect on session length, produced small but statistically significant changes to negative-affect and intentions to return to the game, and does not appear to influence the potential for flow experiences or change positive-affect.

Players could choose when to stop in the control and fatigue conditions only. We labelled each participant based on the recurring themes of satisfaction, discomfort, and reasons external to the game (see Section 5.7 for details). A minority of players stopped because they were satisfied. For the fatigue condition, most players stopped because of discomfort and half of those explicitly mentioned fatigue mechanics as part of their reason.

6.2 Contributions

This dissertation contributes to knowledge by offering subjective user-experience approach to evaluating session length policies. To the best of our knowledge, it is the

[32] Based on suggestions given by Cohen (1988) and Miles & Shevlin (2001) as given by (http://imaging.mrc-cbu.cam.ac.uk/statswiki/FAQ/effectSize)

first text to do so. It also acts to consolidate and supplement the findings presented in Davies and Blake (2016). The user-experience (UXG) of digital games manifests through a variety of game dynamics including those of the player and the context in which they play, and design elements unique to a game and those features common to its genre. We suggest that session length policies may change the fundamental player experience from positive to negative if not carefully applied. Our results inform as to whether the popular implementations of behaviour and design policies do change the UXG in this manner, and we discuss if these attempts are worth the effort required. Our findings contradict the apparent notion that design policies necessarily reduce session length, although we argue that they may assist in reducing overall screen time.

6.2.1 Behaviour policies

We have demonstrated that behaviour policies, such as our shutdown condition, do change a player's session length when the allocated period is shorter than normal. Session length policies intend to reduce playtime, and to that end, our implementation is successful. We identified only one consistent change to UXG in our experiment: a small increase negative-affect, particularly in the *upset* dimension. This result supports the policy because apart from feeling a little *upset*, players still have positive user-experiences when the policies are in effect.

We identified that players subject to behaviour policies wish to continue playing sooner than they did when exposed to fatigue mechanics of design policies. Although we expected this observation, given the stories of players using other national identity documents to continue (see Section 1.2.3), it is no less concerning. The shutdown policy ostensibly presumes that all digital games are detrimental and enforces periodic abstinence to prevent excessive use. In doing so, the policy disregards the positive outcomes associated with longer sessions, such as sustained online social interaction when face-to-face interaction is not possible. Furthermore, it does not consider that harm may be done through the practice of abstinence, such as increasing salient game related thoughts, using another's identity to play, or other game seeking behaviours that potentially break minor laws.

It is these types of behaviours that strengthen the association between addiction and digital games: take a risk to get what the player wants. However, we suspect another reason for the why players go to great lengths for their activity. We briefly discussed in Section 1.2.3 the following scenario: when players lose control over their gaming behaviour, and play time becomes a scarce resource, they may become subject to psychological reactance (Brehm & Brehm, 1981) [33] and overvalue the activity. Whatever the reason, this phenomenon requires further study. At the very least,

[33] A state of mind where low-value opportunities seem more valuable to people when that opportunity's availability is limited.

authorities that wish to employ behaviour policies should be aware of potential side effects such as the psychological reactance, especially in adolescents.

Large-scale implementations, such as those seen in South Korea and Taiwan, include legislation and enforcement, which require time, money, and people to ensure effective execution. Thus, the benefits to their respective societies ought to warrant those costs. Given the reportedly low incidence of problems, it is possible that behaviour policies are not strictly necessary in all circles of society.

This type of policy is a viable option in smaller contexts such as the home environment. Even if games are strictly benevolent, youths stand to benefit from non-gaming activities during their formative years. Before they master self-control may also need external moderation of the activities that they enjoy the most, and especially if those activities are fun. The shutdown system is simple and scalable in that it does not require intricate knowledge of game mechanics nor extensive modifications to the game itself. It follows that parents, guardians, or partners can manage session length by discussion, or with trivial changes to a script like our own if they possess the inclination. Furthermore, there is no indication that the system degrades player experience, at least not significantly.

That being said, the policy is a reaction to the moral panic (see Section 1.1; Bowman, 2015) inspired by video games. In the attempt to mitigate the harms purportedly attributed to video games, the policy does little to consider the complexity of the medium. It may allow players the opportunity to experience the content of a game, but the strict nature denies a player the freedom to use the medium in a way that best benefits them. The policy is ignorant about the effects of agency removal associated with forced hiatus. Players will become upset, and react in a way that restores the agency denied to them by the policy. Our correspondence with participants suggests that the session may end before game conflict and player excitement have resolved. Policy enforcers would need to find ways to avoid these situations, by allowing resolution to occur. We cannot recommend it as a suitable general strategy for minimizing the risks associated with digital games – at least not in the form presented here.

Successful implementation of behaviour policy requires that the enforcer understands how game design affects the player's motivations and goals including their intended session length. Enforcers ought to consider how individual player and context variables affect the outcome of individual play session. We recommend that systems give players sufficient warning of the impending shutdown, or utilize a dynamic but limited period that shuts down only when players have reached a suitable exit point. The latter suggestion is more lenient in that session length is not absolute. This idea resembles the fatigue system in philosophy. Both recommendations allow players to reach a suitable exit point before shutting down the game, but require more work by enforcers for large-scale implementation, and certainly require research to explore their actual effects.

6.2.2 Design polices

We were surprised to see that the fatigue condition consistently increases session length. Our data shows no consequential changes to the measured aspects of subjective UX in games despite enforcing significant restrictions on the pervasive and prominent loot mechanic. Although these findings seemingly are unexpected, because they contravene the policy maker's intention to reduce screen time, we do not conclude that it is a negative outcome. Whereas the system fails to reduce session length, it may in fact moderate hiatus length: we demonstrated a decrease in game-seeking intentions relative to that of our own shutdown condition. If players behave as they intend to, then the fatigue system reduces cumulative screen time without negatively affecting the overall user experience at all.

The fatigue system ostensibly recognizes that frequent marathon sessions are problematic, but that players react negatively when they lose agency. Our data shows that when players choose their own hiatus point, the experience less negative-affect and game-seeking intentions caused by the shutdown condition. Design policies additionally allows high flow, strong positive emotions and weak negative emotions. Participants consistently played for longer under this policy than any other.

Each game has unique mechanics that requires a bespoke implementation of the fatigue elements. This requires work by the game's developers, or by third party developers with extensive knowledge of the game and access to the game's code through modification tools. Implementation requires technical work, and that developers or service providers agree to the terms set by policy makers. This is a problem with design policies in general, as developers must implement the fatigue system in a way that best suits a game's dynamics. If the player switches to another game before the fatigue begins, there is no guarantee that session length can be monitored across games, or that the fatigue mechanics will produce the same effect. Digital game distribution software, such as Steam, can monitor total game time if players must launch their games through the Steam application, but this means that the fatigue systems must be able to interpret the information provided by the application and not rely on their own measures. Furthermore, a player may choose to play a loot free genre. A fatigue system cannot hope to influence players in these games. Thus, such policies are viable on a large scale only, where resources such as designers, developers, time, and money are available to create bespoke fatigue systems for each game.

6.2.3 Policies compared: which is better?

Digital games are popular (see Section 1), and that most players stand to benefit from the hobby (see Section 2.2). It is also clear that there are some risks associated with digital games that should not be ignored (see Section 2.2.2). The policies addressed in this dissertation are designed to mitigate these risks, but are not without their own problems (see Section 1.1). We have shown that neither the behaviour policies nor design policies hinder the medium's ability to produce positive and meaningful experiences. The behaviour policy succeeds in reducing session length where the design policy fails, but this does not mean that the risks disappear with a behaviour policy in place. Recall that games are a complex entertainment medium. The experiences players have been functions of design (genre and game) and behaviours (player and context) (see Section 2.2).

When compared to our control condition, our experimental systems do not appear to degrade positive subjective user experiences measured here: *flow* and *positive-affect*. In fact, our participants appear to experience more intense flow states and comparable positive-affect in our activity than people do in variety of other activities (see Section 5.4 and Section 5.5.1). These findings are encouraging because both policies appear to maintain the beneficial dynamics that gaming provides. From here our discussion comparing these policies can focus on where each might be applied, rather than accepting one while rejecting the other.

The behaviour policy, as presented here, generates positive user experiences, especially moderate to strong flow states, and maintains relatively high positive-affect. It is not without some problematic dynamics. It does not consider genre, game, or player, and considers context insofar as to assume that extended screen time is detrimental before becoming excessive. The policy appears to breed reactance in players by denying them agency, and ignoring much of modern online gaming's charm: the freedom to shed one's real world persona and act in different virtual socially interactive environments.

The design policy presented here does consider genre and game. It subverts those design aspects, such as rewards, that motivate continued play. It also acknowledges the player variable, facilitating agency by allowing players the freedom to play as and when they choose. Furthermore, it is sensitive to the context of play. In no way does the policy force players into situations that may tarnish their social capital by allowing players to continue when their peers rely on them. Our results show that the policy mitigates the reactance we saw with the behaviour policy. The fatigue policy may increase player satisfaction in any given session which leads to the observed decreases in the desire to return to the game.

The policy does increases session length, and there is no guarantee that players will end a session because of it. The restrictive element may be required for those players that struggle to control their own behaviour. It stands to reason that the best policy for mitigating the risks associated with digital games will rely on the interplay of design and behaviour policies.

6.3 Shortcomings and suggested research

User-experience research is complex because of the many factors that define any given session, including genre, game, player and context. Our experiment sought to control as many variables as possible but in doing so we introduced dynamics that changed the validity of our results in a variety of ways. Most notably, the laboratory conditions lowered ecological validity. Here we discuss the key dynamics that likely affect the applicability of our findings.

Not a treatment

As with most debates, commentators and researchers have some stake in the topic and tend to argue passionately for or passionately against the idea that games cause problems for players. Our reading of digital games literature exposes both positive and negative effects, though it appears that only a minority of players suffer adverse effects. It remains unclear if digital games cause these problems. Our results do not contribute to this argument, nor did we intend for them to do so. Furthermore, the policies should not be considered a form of treatment for problematic behaviours. It is also evident that digital games benefit players, but their prevalence and whether they differ significantly from other stimulating hobbies remains to be seen. Session length continues to be policed and controlled while the causal effects of problematic behaviours remain unclear. Our data cannot inform the clinical effectiveness of such policies.

No social interaction

Our chosen genre, an action Role-Playing Game (RPG), represents those games commonly related to problematic gaming in the broader games discourse: Massively Multiplayer Online (MMO) RPGs. It contains similar aesthetics and mechanics, and appeals to a similar type of player. However, the game (and our experiment's design) lack the social interactions and influences associated with massively multiplayer games; aspects that do contribute to player choices, particularly with respect to hiatus and to session length. Many of our participants chose to end their session because of a social pressure to stop playing: transport, meetings, or class performance for example. Our experimental system specifically lacked any social pressure to keep playing: encouragement from party members to continue, time critical matches or dungeon raids. That being said, there is evidence to suggest that even in MMOs, players spend only a third of their time engaging with other party members. It is possible that our findings do reflect common behaviours in these games.

One session does not equal one game

Our experiment presumes that players will only play one game in each session. In practice, players may play many games and seek different experiences during a single session. A person may begin playing an MMORPG, but switch to an FPS mere minutes before the fatigue mechanics instantiates, and continue playing the FPS for many hours. As such we cannot generalize our results to all sessions, but only to those in which players play a single game. Further research is required to explore player behaviours and user-experience when playing multiple games in a single session.

Psychological reactance

Our own conclusions regarding the behaviour policies are problematic. Behaviour policies turn session length into a scarce resource (see Section 1.2.3). We speculate that players may place more value on the time allocated to them each day, and use more than they ordinarily would simply because the limit exists. Cialdini (2007) discusses some human-behavioural phenomena related to the *scarcity* of opportunities using *psychological reactance* theory (Brehm & Brehm, 1981); people attach far more value to opportunities when they have limited availability. We think psychological reactance may be a useful paradigm with which to investigate the shutdown condition; it may explain player intentions to play again sooner than they did after the control and fatigue conditions, and to some extent the identity fraud phenomenon discussed in Section 1.2.3. That discussion remains speculative in this dissertation because we found the theory after designing the experiment and do not wish to form conclusions based upon a post-hoc rationalization. Therefore, we strongly suggest further work, exploring psychological reactance as a model for game seeking behaviours.

No end game data

Our participants experienced only the early-to-mid-game content and mechanics in the experimental system using Torchlight II. Runic Games designed Torchlight II to motivate continued play through story, quests, and loot. Although we have anecdotal evidence suggesting that some players valued loot (see Section 5.3.2) we have no data describing the late-to-end-game contexts wherein players have witnessed the story's conclusion and completed most, if not all, in-game quests. Here character progression is slow by design, requiring exponentially more XP and time to reach new levels and the players' primary motivation for continued play may be seeking the best loot, known as legendary items. Our data, therefore, does not represent those players who spend the most time in Torchlight II. However, Torchlight II completion statistics show that of the three million people who own Torchlight II only fourteen percent complete the story and related quests (Steamspy, 2017). Only seven percent find a legendary item (Steam, 2017). Considering these numbers, we feel that our data represents most players.

Control issues

Table 6.2 Policy-exposure order. The three participant groups, and the order in which they experienced each policy.

Participant group	First session	Second session	Third session
Sequence SCF	Shutdown	Control	Fatigue
Sequence FSC	Fatigue	Shutdown	Control
Sequence CFS	Control	Fatigue	Shutdown

Our design intended to perform direct comparisons of the policies. To this end, we used a crossover experiment design, which resulted in two noteworthy shortcomings. Firstly, we recognize that our control condition is neither a policy nor a true representation of a so-called normal gameplay experience due to the laboratory context. However, it does offer players both agency, and an unmodified gaming experience that is sufficiently different from the experimental conditions. Secondly, our methods do not exhaustively control for order and sampling effects. We employed a Latin square counterbalance, which partially counterbalances exposure to the policies in one of three sequences (see Table 6.2). This method ensures that each condition occurs at each session so that no single condition affects subsequent sessions too much. However, by using this method it is more common for some conditions to precede others than vice versa. For example, the shutdown condition precedes the control condition in sequence *SCF* and sequence *FSC*, whereas the control condition precedes the shutdown condition once only in sequence *CFS*. This means that the experience of shutdown condition potentially affects the experience of the control twice as often. The same occurs for each pair of conditions. Our results do not indicate that this confounded our data as we found no consequential carry-over effects of this nature. A full counterbalance utilizing six unique participant groups would not have introduced more error to our analysis, and would strengthen internal validity. We suggest that further research make use of it.

Neither policy described in Section 1.2 are ‘secret’ implementations; the policies are intended to be a permanent addition to gaming practice and players know they exist. We attempted to maintain ecological validity by informing participants about the policy, and how much time they had to play. It is likely that this produced reactivity in the participants who adjusted their behaviour accordingly. While participant *F* did so in a way that is of qualitative interest to us, others may playing the role of good subject and brought their behaviours in line with what they thought the experimenter wanted to see (Nichols & Maner, 2008). Still others may have resented the imposition placed upon them by the experiment, and played for longer to spite the policy (Masling, 1966). Future research should take care to minimize the influence of these effects. Using a single or double-blind design should control for these types of reactivity. However, this design would only test for the effects of the deception, or the utility of a design policy.

Design policy: handicap or promote?

The laboratory nature of the experiment limited the amount of time that participants were able or willing to spend in the study. Our data represents limited exposure to each policy. Players experience one policy repeatedly in their context, and continued exposure may breed frustration not seen in our sample. Rob Pardo’s talk about fatigue systems in *World of Warcraft* suggest that their beta-testers[34] became frustrated when Blizzard framed the fatigue mechanics as a handicap, but beta-testers enjoyed the mechanics when they framed normal gameplay as a boost, and the fatigued mechanics as normal gameplay (Pardo, 2010). A longer study is required to verify if this is a common effect.

Mixed methods designs

Large scale or longitudinal studies appear to address the shortcomings mentioned above. However, many of the variables addressed by our study may confound the results, or obfuscate the cause of effects. Therefore, we suggest that large-scale longitudinal studies must supplement small-scale laboratory studies such as ours to gain the most understanding of player behaviours.

---

[34] Select members of the public who play a game to identify issues that developers have missed before the game is released to the general public

Player types

Our research does only consider those people who play RPGs currently or sometime in the past; it does not use a more complex player typology to frame our experiment or inform our results. Player type defines player motivations by how they choose to act (Bartle, 1996), how they feel (Lazzaro, 2004), the extent to which they commit (Y. Poels, Annema, Verstraete, Zaman, & De Grooff, 2012) and how they relate to their avatar, all of which affect player decision making in turn. Although our experiment uses the cross-over elements to minimize the effects of these differences study, it does not collect data in a form that allows us to identify these types and relationships, and explain our observations from a typological perspective. Our sample ostensibly reflects the experiences of a range of players, potentially normalizing our results. Any further study stands to benefit from considering only certain player types or players with relevant personality factors. More focused research may facilitate more nuanced discussion that addresses why our sample deviated from the expected results. Given the low incidence of problematic gameplay, our sample may not include those who stand to benefit the most from these policies. Thus, their reactions to the policies are unaccounted for.

Appropriate instruments

Our operational definition for user-experience in games describes those aspects important to our research questions. Both instruments used are valid and reliable for measuring flow and state affect however, they may not be perfectly suited for playing games as an activity. The International Positive And Negative Affect Schedule Short Form is short and excludes some emotions commonly associated with digital games, for example those identified by Lazzaro (2004). Furthermore, our definition of UX in games is not exhaustive. Further research stands to benefit from using instruments designed for use in video games, such as the Game Experience Questionnaire (Ijsselsteijn et al., 2013), particularly in light of our discovering the somewhat hidden validation report (K. Poels et al., 2007a; see Section 2.5)

6.4 Conclusion

As video game researchers, we must acknowledge that our field is complicated; many variables influence the types of effects we observe, and conclusive evidence is rare especially when generalizing to all players and games. The policies we investigated in the present study do not appear to degrade the positive experiences associated games, particularly those that contribute to lasting well-being. Although our findings show benign results, we cannot ultimately conclude that these policies are a desirable institution. We acknowledge that policy makers try to control session length because they are socially accountable to their people, or have concern for the well-being of the gamers in their custody. That they should seek to control their population’s gaming behaviours to possibly prevent a rare and uncertain psychopathology is itself problematic; policy treats a sometimes symptom as if it were a pathology. Policy makers and enforcers, and game developers would do well to use our findings not as confirmation that their systems are working, but to identify potentially unforeseen side effects of their good intentions. Behaviour policy requires human resources to ensure enforcement, more so if players resist and circumvent the systems. Design policy implementation requires development time and financial support; games cannot generate revenue if players take avoid the game because of policy implementation. Our contributions can assist game designers should they be required to implement these policies. In the interest of retaining their player base, they will know to grant players as much agency as possible when ending sessions. They will know that fatiguing in loot mechanics do not necessarily degrade the positive experiences that they are crafting for their players.

7 REFERENCES

Aarseth, E., Bean, A. M., Boonen, H., Colder Carras, M., Coulson, M., Das, D., … Van Rooij, A. J. (2016). Scholars' open debate paper on the World Health Organization ICD-11 Gaming Disorder proposal. *Journal of Behavioral Addictions*, *6*(3), 1–4. https://doi.org/10.1556/2006.5.2016.088

Abrams, S. (2010). Breaking News: China Gamers Support Real-ID System. Retrieved February 23, 2017, from http://www.chinahearsay.com/breaking-news-china-gamers-support-real-id-system/

Adachi, P. J. C., & Willoughby, T. (2013). Do Video Games Promote Positive Youth Development? *Journal of Adolescent Research*, *28*(2), 155–165. https://doi.org/10.1177/0743558412464522

Adams, E. W. (2004). Postmodernism and the Three Types of Immersion. Retrieved January 13, 2017, from https://www.gamasutra.com/view/feature/130531/the_designers_notebook_.php

Allahverdipour, H., Bazargan, M., Farhadinasab, A., & Moeini, B. (2010). Correlates of video games playing among adolescents in an Islamic country. *BMC Public Health*, *10*(1), 286. https://doi.org/10.1186/1471-2458-10-286

Allaire, J. C., McLaughlin, A. C., Trujillo, A., Whitlock, L. A., LaPorte, L., & Gandy, M. (2013). Successful aging through digital games: Socioemotional differences between older adult gamers and Non-gamers. *Computers in Human Behavior*, *29*(4), 1302–1306. https://doi.org/10.1016/j.chb.2013.01.014

American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders* (5th ed.). Arlington, VA: American Psychiatric Publishing. https://doi.org/10.1176/appi.books.9780890425596

Arnason, S. L. (2016). *Regulating Online Games in China: Policy, Practice, Innovation, and Change. PhD – University of Edinburgh*. The University of Edinburgh.

Retrieved from http://hdl.handle.net/1842/17032

Banks, J. (2013). *Human-Technology Relationality and Self-Network Organization: Players and Avatars in World of Warcraft*. Colorado State University.

Banks, J., & Bowman, N. D. (2016). Avatars are (sometimes) people too: Linguistic indicators of parasocial and social ties in player-avatar relationships. *New Media & Society*, *18*(7), 1461444814554898-. https://doi.org/10.1177/1461444814554898

Bartle, R. (1996). Hearts, Clubs, Diamonds, Spades: Players who suit MUDs. *Journal of MUD Research*, *6*(1), 19. https://doi.org/10.1007/s00256-004-0875-6

Bean, A. M., Ferro, L. S., Vissoci, J. R. N., Rivero, T., & Groth-Marnat, G. (2016). The emerging adolescent World of Warcraft video gamer: A five factor exploratory profile model. *Entertainment Computing*, *17*, 45–54. https://doi.org/10.1016/j.entcom.2016.08.006

Bean, A. M., & Groth-Marnat, G. (2016). Video gamers and personality: A five-factor model to understand game playing style. *Psychology of Popular Media Culture*, *5*(1), 27–38. https://doi.org/10.1037/ppm0000025

Bethesda Game Studios. (2011). Skyrim [Video Game]. Bethesda Softworks. Retrieved from https://elderscrolls.bethesda.net/en/skyrim?

Bhaumik, D. K., Roy, A., Aryal, S., Hur, K., Duan, N., Normand, S.-L. T., … Gibbons, R. D. (2008). Sample Size Determination for Studies with Repeated Continuous Outcomes. *Psychiatric Annals*, *38*(12), 765–771. https://doi.org/10.3928/00485713-20081201-01

Blizzard Entertainment. (1998). StarCraft. Blizzard Entertainment. Retrieved from https://starcraft.com/en-us/

Blizzard Entertainment. (2004). World of Warcraft. Retrieved September 7, 2017, from https://worldofwarcraft.com/en-us/

Bo, Q. (2010). Gamers back real-name registration: Poll. Retrieved February 23, 2017, from http://www.chinadaily.com.cn/china/2010-08/18/content_11167094.htm

Bowman, N. D. (2016). The rise (and refinement) of moral panic. In R. Kowert & T. Quandt (Eds.), *The Video Game Debate: Unraveling the physical, social, and psychological effects of digital games* (pp. 22–38). New York: Routledge.

Bowman, N. D., Weber, R., Tamborini, R., & Sherry, J. (2013). Facilitating Game Play: How Others Affect Performance at and Enjoyment of Video Games. *Media Psychology*, *16*(1), 39–64. https://doi.org/10.1080/15213269.2012.742360

Brehm, S. S., & Brehm, J. W. (1981). *Psychological reactance: A theory of freedom and control*. New York: Academic Press.

Brockmyer, J. H., Fox, C. M., Curtiss, K. A., Mcbroom, E., Burkhart, K. M., & Pidruzny, J. N. (2009). The development of the Game Engagement Questionnaire: A measure of engagement in video game-playing. *Journal of Experimental Social Psychology*, *45*(4), 624–634. https://doi.org/10.1016/j.jesp.2009.02.016

Brown, E., & Cairns, P. (2004). A grounded investigation of game immersion. In *CHI '04 Extended Abstracts on Human Factors in Computing Systems* (pp. 1297–1300).

Vienna: ACM. https://doi.org/10.1145/985921.986048

Burge, J. (2009). *Petric Verdict II; Video games and the brain [Video file]*. YouTube. Retrieved from https://www.youtube.com/watch?v=WY0f0DzfdIw

Caillois, R. (1961). *Man, Play, and Games*. Urbana and Chicago: University of Illinois Press.

Cairns, P., Cox, A., & Nordin, A. I. (2014). Immersion in Digital Games: Review of Gaming Experience Research. In M. C. Angelides & H. Agius (Eds.), *Handbook of Digital Games* (pp. 337–361). Hoboken, NJ, USA: John Wiley & Sons, Inc. https://doi.org/10.1002/9781118796443.ch12

Calvillo-Gámez, E. H., Cairns, P., & Cox, A. L. (2010). Assessing the Core Elements of the Gaming Experience. In R. Bernhaupt (Ed.), *Evaluating User Experience in Games* (pp. 47–71). London: Springer-Verlag. https://doi.org/10.1007/978-1-84882-963-3_4

Capcom. (1992). Street Fighter II. Capcom. Retrieved from https://streetfighter.com/

Chen, J. (2006). *Flow in Games*. *jenovachen.com*. University of Southern California. Retrieved from http://www.jenovachen.com/flowingames/Flow_in_games_final.pdf

Chen, J. (2007). Flow in games (and everything else). *Communications of the ACM*, *50*(4), 31–34. https://doi.org/10.1145/1232743.1232769

China Daily. (2007). The more they play, the more they lose. [archive]. Retrieved March 11, 2018, from http://en.people.cn/200704/10/eng20070410_364977.html

Chory, R. M., & Goodboy, A. K. (2011). Is Basic Personality Related to Violent and Non-Violent Video Game Play and Preferences? *Cyberpsychology, Behavior, and Social Networking*, *14*(4), 191–198. https://doi.org/10.1089/cyber.2010.0076

Chou, T.-J., & Ting, C.-C. (2004). The role of flow experience in cyber-game addiction. *Cyberpsychology & Behavior*, *6*(6), 663–675. https://doi.org/10.1089/109493103322725469

Cialdini, R. B. (2007). *Influence: the psychology of persuasion* (First Coll). New York: HaperCollins.

Cochran, W. G., & Cox, G. M. (1992). *Experimental Designs* (2nd ed.). New York: Wiley.

Cohen, J. (1988). Statistical power analysis for the behavioral sciences. *Statistical Power Analysis for the Behavioral Sciences*. Hillsdale, N.J: L. Erlbaum Associates. https://doi.org/10.1016/B978-0-12-179060-8.50012-8

Connolly, T. M., Boyle, E. A., MacArthur, E., Hainey, T., & Boyle, J. M. (2012). A systematic literature review of empirical evidence on computer games and serious games. *Computers & Education*, *59*(2), 661–686. https://doi.org/10.1016/j.compedu.2012.03.004

Coulson, M., & Ferguson, C. J. (2016). The Influence of Digital Games on Agression and Violent Crime. In R. Kowert & T. Quandt (Eds.), *The Video Game Debate: Unraveling the physical, social, and psychological effects of digital games* (pp. 54–

73). New York: Routledge.

Cowley, B., Charles, D., Black, M., & Hickey, R. (2008). Toward an understanding of flow in video games. *Computers in Entertainment*, *6*(2), 20:1-27. https://doi.org/10.1145/1371216.1371223

Csikszentmihalyi, M. (1975). Play and Intrinsic Rewards. *Journal of Humanistic Psychology*, *15*(3), 41–63. https://doi.org/10.1177/002216787501500306

Csikszentmihalyi, M. (1988). The flow experience and its significance for human psychology. In M. Csikszentmihalyi & I. S. Csikszentmihalyi (Eds.), *Optimal experience: Psychological studies of flow in consciousness* (pp. 15–35). Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511621956.002

Csikszentmihalyi, M. (1990). *Flow: The Psychology of Optimal Experience* (1st ed.). New York: Harper & Row.

Csikszentmihalyi, M., & LeFevre, J. (1989). Optimal Experience in Work and Leisure. *Journal of Personality and Social Psychology*, *56*(5), 815–822. https://doi.org/10.1037/0022-3514.56.5.815

Csikszentmihalyi, M., & Massimini, F. (1985). On the psychological selection of bio-cultural information. *New Ideas in Psychology*, *3*(2), 115–138. https://doi.org/10.1016/0732-118X(85)90002-9

Davies, B., & Blake, E. (2016, March). Evaluating existing strategies to limit video game playing time. *IEEE Computer Graphics and Applications*, *36*(2), 47–57. https://doi.org/10.1109/MCG.2016.25

de Kort, Y., Poels, K., & Ijsselsteijn, W. (2007). Digital Games as Social Presence Technology : Development of the Social Presence in Gaming Questionnaire (SPGQ). In L. Moreno (Ed.), *PRESENCE 2007: The 10th Annual International Wokshop on Presence* (pp. 195–203). Barcelona: Starlab Barcelona S.L. Retrieved from http://home.ieis.tue.nl/ydkort/de kort et al Digital games as social presence technology PRESENCE 2007.pdf

DeKoven, B. (n.d.). Dekoven - Of fun and flow. Retrieved from http://www.deepfun.com/funflow.htm

Dennaton Games. (2012). Hotline Miami [Video Game]. Devolver Digital.

Dennie, T. M. (2012). *Perception of Autonomy and its Effect on Intrinsic Motivation, Immersion, and Performance*. Western Carolina University. Retrieved from https://search.proquest.com/docview/1032561950?accountid=14500

Dontnod Entertainment. (2015). Life is Strange [Video Game]. Square Enix.

Eisenberger, R., Jones, J. R., Stinglhamber, F., Shanock, L., & Randall, A. T. (2005). Flow experiences at work: For high need achievers alone? *Journal of Organizational Behavior*, *26*(7), 755–775. https://doi.org/10.1002/job.337

Elson, M., & Ferguson, C. J. (2013). Gun violence and media effects: challenges for science and public policy. *British Journal of Psychiatry*, *203*(5), 322–324. https://doi.org/10.1192/bjp.bp.113.128652

Engeser, S. (2012a). Appendix: Flow Short Scale. In S. Engeser (Ed.), *Advances in Flow*

Research (First, pp. 201–202). New York, NY: Springer. https://doi.org/10.1007/978-1-4614-2359-1

Engeser, S. (2012b). Theoretical integration and future lines of flow research. In S. Engeser (Ed.), *Advances in Flow Research* (pp. 187–199). Ney York, NY: Springer. https://doi.org/10.1007/978-1-4614-2359-1_10

Engeser, S., & Rheinberg, F. (2008). Flow, moderators of challenge-skill-balance and performance. *Motivation and Emotion*, *32*(3), 158–172. https://doi.org/10.1007/s11031-008-9102-4

Engeser, S., & Schiepe-Tiska, A. (2012). Historical lines and an overview of current research on flow. In S. Engeser (Ed.), *Advances in Flow Research* (pp. 1–22). New York, NY: Springer. https://doi.org/10.1007/978-1-4614-2359-1_1

Entertainment Software Association. (2016). *Essential Facts about the computer and video game industry*. *Entertainment Software Association*.

Ermi, L., & Mäyrä, F. (2007). Fundamental Components of the Gameplay Experience: Analysing Immersion. In J. Jenson & S. De Castell (Eds.), *Worlds in Play: international perspectives on digital games research* (Paperback, pp. 37–55). New York, NY: Peter Lang.

Ernkvist, M., & Ström, P. (2008). Enmeshed in Games with the Government: Governmental Policies and the Development of the Chinese Online Game Industry. *Games and Culture*, *3*(1), 98–126. https://doi.org/10.1177/1555412007309527

Falconer, C. J., Slater, M., Rovira, A., King, J. A., Gilbert, P., Antley, A., & Brewin, C. R. (2014). Embodying Compassion: A Virtual Reality Paradigm for Overcoming Excessive Self-Criticism. *PLoS ONE*, *9*(11). https://doi.org/10.1371/journal.pone.0111933

Fang, X., & Zhao, F. (2010). Personality and enjoyment of computer game play. *Computers in Industry*, *61*(4), 342–349. https://doi.org/10.1016/j.compind.2009.12.005

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G * Power 3 : A flexible statistical power analysis program for the social , behavioral , and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/bf03193146

Ferguson, C. J. (2007). The good, the bad and the ugly: A meta-analytic review of positive and negative effects of violent video games. *Psychiatric Quarterly*, *78*(4), 309–316. https://doi.org/10.1007/s11126-007-9056-9

Ferguson, C. J. (2015). Do Angry Birds Make for Angry Children? A Meta-Analysis of Video Game Influences on Children's and Adolescents' Aggression, Mental Health, Prosocial Behavior, and Academic Performance. *Perspectives on Psychological Science*, *10*(5), 646–666. https://doi.org/10.1177/1745691615592234

Ferguson, C. J. (2016). New American Academy Of Pediatrics Screen Time Recommendations Still Don't Make a Passing Grade. Retrieved February 6, 2017, from http://www.huffingtonpost.com/entry/5814a3fae4b08301d33e0a16?timestamp=1477919471821

Fleming, R. (2011). Man attacks child over Call of Duty | Digital Trends. Retrieved February 2, 2017, from http://www.digitaltrends.com/gaming/man-attacks-child-over-call-of-duty/

Foster, E. M., & Watkins, S. (2010). The Value of Reanalysis : TV Viewing and Attention Problems Author (s): E . Michael Foster and Stephanie Watkins Published by : Wiley on behalf of the Society for Research in Child Development Stable URL : http://www.jstor.org/stable/40598984 REFERENCES, *81*(1), 368–375. Retrieved from https://doi.org/10.1111/j.1467-8624.2009.01400.x

Frasca, G. (1999). Ludology Meets Narratology: Similitude and differences between (video) games and Narrative. Retrieved October 6, 2017, from http://www.ludology.org/articles/ludology.htm

Fredenburg, H. A., & Silverman, M. J. (2014). Effects of music therapy on positive and negative affect and pain with hospitalized patients recovering from a blood and marrow transplant: A randomized effectiveness study. *The Arts in Psychotherapy*, *41*(2), 174–180. https://doi.org/10.1016/j.aip.2014.01.007

Fredrickson, B. L. (2013). Positive Emotions Broaden and Build. In P. Devine & P. Ashby (Eds.), *Advances in Experimental Social Psychology* (1st ed., Vol. 47, pp. 1–53). San Diego: Academic Press. https://doi.org/10.1016/B978-0-12-407236-7.00001-2

Frictional Games. (2010). Amnesia: The Dark Descent [Video Game]. Frictional Games. Retrieved from http://www.amnesiagame.com/

Fu, F.-L., Su, R.-C., & Yu, S.-C. (2009). EGameFlow: A scale to measure learners' enjoyment of e-learning games. *Computers and Education*, *52*(1), 101–112. https://doi.org/10.1016/j.compedu.2008.07.004

Games, L. (1990). Monkey Island [Video Game]. Lucasfilm Games. Retrieved from http://monkeyisland.wikia.com/wiki/Monkey_Island_series

Gerling, K. M., Birk, M., Mandryk, R. L., & Doucette, A. (2013). The Effects of Graphical Fidelity on Player Experience. In *Proceedings of International Conference on Making Sense of Converging Media* (p. 229:236). New York, NY, USA: ACM. https://doi.org/10.1145/2523429.2523473

Girden, E. R. (1992). *ANOVA: Repeated Measures*. Newbury Cliff, CA: Sage Publications.

Granic, I., Lobel, A., & Engels, R. C. M. E. (2014). The benefits of playing video games. *American Psychologist*, *69*(1), 66–78. https://doi.org/10.1037/a0034857

Gray, J. A. (1970). The psychophysiological basis of introversion-extraversion. *Behaviour Research and Therapy*, *8*(3), 249–266. https://doi.org/10.1016/0005-7967(70)90069-0

Gray, J. A., & McNaughton, N. (2003). *The neuropsychology of anxiety : an enquiry into the functions of the septo-hippocampal system* (2nd ed.). Oxford: Oxford University Press.

Green, C. S., & Bavelier, D. (2012). Learning, attentional control, and action video games. *Current Biology*, *22*(6), R197–R206. https://doi.org/10.1016/j.cub.2012.02.012

Green, M. C., Brock, T. C., & Kaufman, G. F. (2004). Understanding media enjoyment: The role of transportation into narrative worlds. *Communication Theory*, *14*(4), 311–327. https://doi.org/10.1111/j.1468-2885.2004.tb00317.x

Griffiths, M. D. (2005). A "components" model of addiction within a biopsychosocial framework. *Journal of Substance Use*, *10*(4), 191–197. https://doi.org/10.1080/14659890500114359

Griffiths, M. D. (2016). Gaming Addiction and Internet Gaming Disorder. In R. Kowert & T. Quandt (Eds.), *The Video Game Debate: Unraveling the physical, social, and psychological effects of digital games* (pp. 74–93). New York: Routledge.

Gunter, W. D., & Daly, K. (2012). Causal or spurious: Using propensity score matching to detangle the relationship between violent video games and violent behavior. *Computers in Human Behavior*, *28*(4), 1348–1355. https://doi.org/10.1016/j.chb.2012.02.020

Hamari, J., & Tuunanen, J. (2014). Player Types : A Meta-synthesis. *Selected Articles from the DiGRA Nording 2012 Conference: Local and Global-Games in Culture and Society*, 29–53. https://doi.org/10.1111/j.1083-6101.2006.00301.x

Hartmann, T., Wirth, W., Schramm, H., Klimmt, C., Vorderer, P., Gysbers, A., … Sacau, A. M. (2016). The spatial presence experience scale (SPES): A short self-report measure for diverse media settings. *Journal of Media Psychology*, *28*(1), 1–15. https://doi.org/10.1027/1864-1105/a000137

Hektner, J. M., & Csikszentmihalyi, M. (1996). *A Longitudinal Exploration of Flow and Intrinsic Motivation in Adolescents*. *Educational Resources Information Center (ERIC)*. New York.

Heo, J., Oh, J., Subramanian, S. V, Kim, Y., & Kawachi, I. (2014). Addictive internet use among Korean adolescents: A national survey. *PLoS ONE*, *9*(2), 1–8. https://doi.org/10.1371/journal.pone.0087819

Hoenig, J. M., & Heisey, D. M. (2001). The abuse of power. *The American Statistician*, *55*(1), 19–24. https://doi.org/10.1198/000313001300339897

Hoffman, H. G., Chambers, G. T., Meyer, W. J., Arceneaux, L. L., Russell, W. J., Seibel, E. J., … Patterson, D. R. (2011). Virtual reality as an adjunctive non-pharmacologic analgesic for acute burn pain during medical procedures. In *Annals of Behavioral Medicine* (Vol. 41, pp. 183–191). https://doi.org/10.1007/s12160-010-9248-7

Huizinga, J. (1949). *Homo Ludens: A Study of the Play-Element in Culture* (English). London: Routledge & Kegan Paul. https://doi.org/10.1177/0907568202009004005

Hussain, Z., & Griffiths, M. D. (2008). Gender swapping and socializing in cyberspace: an exploratory study. *CyberPsychology & Behavior*, *11*(1), 47–53. https://doi.org/10.1089/cpb.2007.0020

IBM. (2013). SPSS 22 [Statistics Package]. Armonk, NY: IBM Corporation.

IBM. (2014). SPSS 23 [Statistics Package]. Armonk, NY: IBM Corporation.

Ijsselsteijn, W., de Kort, Y., Jurgelionis, A., & Bellotti, F. (2007). Characterising and Measuring User Experiences in Digital Games. In R. Bernhaupt & M. Tscheligi

(Eds.), *International Conference on Advances in Computer Entertainment Technology (ACE 2007)* (pp. 1–4). Salzburg, Austria. Retrieved from http://repository.tue.nl/661449

Ijsselsteijn, W., de Kort, Y., & Poels, K. (2013). The Game Experience Questionnaire. Eindhoven: Technische Universiteit Eindhoven. Retrieved from https://pure.tue.nl/ws/files/21666907/Game_Experience_Questionnaire_English.pdf

Ivory, J. D. (2016). A brief history of videogames. In R. Kowert & T. Quandt (Eds.), *The Video Game Debate: Unraveling the physical, social, and psychological effects of digital games* (pp. 1–21). New York: Routledge.

Jacobson, D. (2002). On theorizing presence. *Journal of Virtual Environments*, *6*(1), 8. Retrieved from http://www.jimelwood.net/students/tsukuba/english_s/jacobson_2002.doc

Jennett, C., Cox, A. L., Cairns, P., Dhoparee, S., Epps, A., Tijs, T., & Walton, A. (2008). Measuring and defining the experience of immersion in games. *International Journal of Human Computer Studies*, *66*(9), 641–661. https://doi.org/10.1016/j.ijhcs.2008.04.004

John, O. P., Naumann, L. P., & Soto, C. J. (2008). Paradigm Shift to the Integrative Big Five Trait Taxonomy: History, measurement, and conceptual issues. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *Handbook of personality: Theory and Research* (pp. 114–158). New York: Guildford Press. Retrieved from http://www.colby.edu/psychology/labs/personality/publications/John_et_al_2008.pdf

Johnson, D., & Gardner, J. (2010). Personality, Motivation and Video Games. *Proceedings of the 22nd Conference of the Computer-Human Interaction Special Interest Group of Australia on Computer-Human Interaction - OZCHI '10*, 276. https://doi.org/10.1145/1952222.1952281

Jones, M. G. (1998). *Creating Electronic Learning Environments : Games, Flow, and the User Interface. Proceedings of Selected Research and Development Presentations at the National Convention of the Association for Educational Communications and Technology (AECT)*. Retrieved from https://eric.ed.gov/?id=ED423842

Karim, J., Weisz, R., & Rehman, S. U. (2011). International positive and negative affect schedule short-form (I-PANAS-SF): Testing for factorial invariance across cultures. *Procedia-Social and Behavioral Sciences*, *15*, 2016–2022. https://doi.org/10.1016/j.sbspro.2011.04.046

Kato, P. M., Cole, S. W., Bradlyn, A. S., & Pollock, B. H. (2008). A video game improves behavioral outcomes in adolescents and young adults with cancer: a randomized trial. *Pediatrics*, *122*(2), e305–e317. https://doi.org/10.1542/peds.2007-3134

Kelly, V. C. (2014). A Primer of Affect Psychology. Retrieved May 31, 2017, from http://www.tomkins.org/wp-content/uploads/2014/07/Primer_of_Affect_Psychology-Kelly.pdf

Kowert, R. (2016). Social Outcomes. In R. Kowert & T. Quandt (Eds.), *The Video Game Debate: Unraveling the physical, social, and psychological effects of digital games* (pp. 94–115). New York: Routledge.

Kowert, R., & Quandt, T. (2016). No Black and White in Video Game Land! Why We Need to Move Beyond simple Explanations in the Video Game Debate. In R. Kowert & T. Quandt (Eds.), *The Video Game Debate: Unraveling the physical, social, and psychological effects of digital games* (pp. 176–189). New York: Routledge.

Kuczmierczyk, A. R., Walley, P. B., & Calhoun, K. S. (1987). Relaxation Training, In Vivo Exposure and Response-prevention in the Treatment of Compulsive Video-game Playing. *Scandinavian Journal of Behaviour Therapy*, *16*(4), 185–190. https://doi.org/10.1080/16506078709455801

Kuesten, C., Chopra, P., Bi, J., & Meiselman, H. (2014). A global study using PANAS (PA and NA) scales to measure consumer emotions associated with aromas of phytonutrient supplements. *Food Quality and Preference*, *33*, 86–97. https://doi.org/10.1016/j.foodqual.2013.11.004

Kuss, D. J., & Griffiths, M. D. (2012). Internet Gaming Addiction: A Systematic Review of Empirical Research. *International Journal of Mental Health and Addiction*, *10*(2), 278–296. https://doi.org/10.1007/s11469-011-9318-5

Kuss, D. J., Griffiths, M. D., & Pontes, H. M. (2017). Chaos and confusion in DSM-5 diagnosis of Internet Gaming Disorder: Issues, concerns, and recommendations for clarity in the field. *Journal of Behavioral Addictions*, *6*(2), 103–109. https://doi.org/10.1556/2006.5.2016.062

Landhäußer, A., & Keller, J. (2012). Flow and its affective, cognitive, and performance-related consequences. In S. Engeser (Ed.), *Advances in Flow Research* (pp. 65–85). New York, NY: Springer. https://doi.org/10.1007/978-1-4614-2359-1_4

Lazzaro, N. (2004). *Why We Play Games: Four Keys to More Emotion Without Story. White Paper*. Retrieved from http://xeodesign.com/xeodesign_whyweplaygames.pdf

Lieberman, J. D., Solomon, S., Greenberg, J., & McGregor, H. a. (1999). A hot new way to measure aggression. *Aggressive Behavior*, *25*(5), 331–348. https://doi.org/10.1002/(SICI)1098-2337(1999)25:5%3C331::AID-AB2%3E3.0.CO;2-1

LimeSurvey GmbH. (2003). LimeSurvey: the online survey tool. *Online*. Retrieved from https://www.limesurvey.org/

Locker, M. (2015). This Place Just Made it Illegal to Give Kids Too Much Screen Time. Retrieved February 7, 2017, from http://time.com/3682621/this-country-just-made-it-illegal-to-give-kids-too-much-screen-time/

Lombard, M., & Ditton, T. (1997). At the Heart of It All: The Concept of Presence. *Journal of Computer-Mediated Communication*, *3*(2). https://doi.org/10.1111/j.1083-6101.1997.tb00072.x

Lombard, M., & Jones, M. (2007). Identifying the (Tele) Presence Literature. *PsychNology Journal*, *5*(2), 197–206. Retrieved from http://psychnology.org/File/PNJ5(2)/PSYCHNOLOGY_JOURNAL_5_2_FULL.pdf#page=97

Lombard, M., & Jones, M. T. (2015). Defining Presence. In M. Lombard, F. Biocca, J. Freeman, W. IJsselsteijn, & R. J. Schaevitz (Eds.), *Immersed in Media* (pp. 13–34).

Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-10190-3_2

Lorentz, P., Ferguson, C. J., & Schott, G. (2015). Editorial: The experience and benefits of game playing. [Editorial]. *Cyberpsychology: Journal of Psychosocial Research on Cyberspace*, *9*(3). https://doi.org/10.5817/CP2015-3-1

Løvoll, H. S., & Vittersø, J. (2014). Can Balance be Boring? A Critique of the "Challenges Should Match Skills" Hypotheses in Flow Theory. *Social Indicators Research*, *115*(1), 117–136. https://doi.org/10.1007/s11205-012-0211-9

Masling, J. (1966). Role-related behavior of the subject and psychologist and its effects upon psychological data. *Nebraska Symposium on Motivation*, *14*, 67–103. Retrieved from http://psycnet.apa.org/record/1967-14474-001

Metcalf, O., & Pammer, K. (2014). Impulsivity and related neuropsychological features in regular and addictive first person shooter gaming. *Cyberpsychology, Behavior and Social Networking*, *17*(3), 147–152. https://doi.org/10.1089/cyber.2013.0024

Mirza-Babaei, P., Nacke, L. E., Gregory, J., Collins, N., & Fitzpatrick, G. (2013). How does it play better?: exploring user testing and biometric storyboards in games user research. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '13* (pp. 1499–1508). https://doi.org/10.1145/2470654.2466200

Moneta, G. B. (2004). The flow experience across cultures. *Journal of Happiness Studies*, *5*(2), 115–121. https://doi.org/10.1023/B:JOHS.0000035913.65762.b5

Moneta, G. B. (2012). On the Measurement and Conceptualization of Flow. In S. Engeser (Ed.), *Advances in Flow Research* (pp. 23–50). New York, NY: Springer. https://doi.org/10.1007/978-1-4614-2359-1_2

Moneta, G. B., & Csikszentmihalyi, M. (1996). The effect of perceived challenges and skills on the quality of subjective experience. *Journal of Personality*, *64*(2), 275–310. https://doi.org/10.1111/j.1467-6494.1996.tb00512.x

Morris, S. B., & DeShon, R. P. (2002). Combining effect size estimates in meta-analysis with repeated measures and independent-groups designs. *Psychological Methods*, *7*(1), 105–125. https://doi.org/10.1037//1082-989X.7.1.105

Murray, J. H. (2005). The Last Word on Ludology v Narratology [Preface to the keynote address]. In *DiGRA 2005 2nd International Conference, "Changing Views: Worlds in Play"* (pp. 1–4). Vancouver. Retrieved from http://www.lcc.gatech.edu/~murray/digra05/lastword.mht

Nacke, L. (2012). Flow in Games : Proposing a Flow Experience Model. In *In Proceedings of the workshop on conceptualising, operationalising and measuring the player experience in videogames at fun and games 2012* (pp. 104–108). Telouse: ACM. Retrieved from https://hcigames.com/download/flow-in-games-proposing-a-flow-experience-model/

Nacke, L., & Lindley, C. (2008). Boredom, Immersion, Flow - A pilot study investigating player experience. In Y. Xiao & E. ten Thij (Eds.), *IADIS international conference gaming 2008: design for engaging experience and social interaction 2008* (pp. 103–107). Amsterdam. Retrieved from http://www.iadisportal.org/digitallibrary/boredom-immersion-flow-—-a-pilot-study-investigating-player-experience

Newzoo. (2014). Asia-Pacific Contributes 82% of the $6Bn Global Games Market Growth in 2014. Retrieved June 19, 2015, from https://newzoo.com/insights/articles/asia-pacific-contributes-82-6bn-global-games-market-growth/

Nichols, A. L., & Maner, J. K. (2008). The good-subject effect: Investigating participant demand characteristics. *Journal of General Psychology*, *135*(2), 151–165. https://doi.org/10.3200/GENP.135.2.151-166

Nordin, A. I., Denisova, A., & Cairns, P. (2014). Too Many Questionnaires: Measuring Player Experience Whilst Playing Digital Games. In M. Luckcuck (Ed.), *Proceedings of the Seventh York Doctoral Symposium on Computer Science & Electronics* (pp. 69–75). York. Retrieved from https://www.cs.york.ac.uk/ftpdir/reports/2014/YCS/494/YCS-2014-494.pdf

Norman, K. L. (2013). GEQ (Game Engagement/experience questionnaire): A review of two papers. *Interacting with Computers*, *25*(4), 278–283. https://doi.org/10.1093/iwc/iwt009

Novak, T. P., & Hoffman, D. L. (1997). Measuring the flow experience among web users. *Interval Research Corporation*, 1–36.

Oliver, M. B., & Bartsch, A. (2010). Appreciation as Audience Response: Exploring Entertainment Gratifications Beyond Hedonism. *Human Communication Research*, *36*(1), 53–81. https://doi.org/10.1111/j.1468-2958.2009.01368.x

Oliver, M. B., Bowman, N. D., Woolley, J. K., Rogers, R., Sherrick, B. I., & Chung, M.-Y. (2016). Video games as meaningful entertainment experiences. *Psychology of Popular Media Culture*, *5*(4), 390–405. https://doi.org/10.1037/ppm0000066

Olson, C. K. (2010). Children's motivations for video game play in the context of normal development. *Review of General Psychology*, *14*(2), 180–187. https://doi.org/10.1037/a0018984

Olson, C. K. (2016). Health Hazards. In R. Kowert & T. Quandt (Eds.), *The Video Game Debate: Unraveling the physical, social, and psychological effects of digital games* (pp. 39–53). New York: Routledge.

Olson, C. K., Kutner, L. A., & Warner, D. E. (2008). The Role of Violent Video Game Content in Adolescent Development: Boys' Perspectives. *Journal of Adolescent Research*, *23*(1), 55–75. https://doi.org/10.1177/0743558407310713

Palmer, M. T. (1995). Interpersonal communication and virtual reality: mediating interpersonal relationships. In F. Biocca & M. R. Levy (Eds.), *Communication in the age of virtual reality* (pp. 277–299). Hillsdale, NJ: L. Erlbaum Associates Inc. Retrieved from http://dl.acm.org/citation.cfm?id=207922.207932

Pardo, R. (2010). GDC Vault - Making a Standard (and Trying to Stick to it!): Blizzard Design Philosophies [Video File]. Retrieved February 7, 2017, from http://www.gdcvault.com/play/1012291/Making-a-Standard-(and-Trying

Parkin, S. (2013). The Video-Game Invasion of Iraq. Retrieved February 4, 2017, from http://www.newyorker.com/tech/elements/the-video-game-invasion-of-iraq

Parkin, S. (2015). *Death by Video Game, Tales of obsession from the virtual frontline - Google Play* (1st ed.). London: Serpent's Tail. Retrieved from https://serpentstail.com/death-by-video-game.html

Pidd, H. (2012). Anders Breivik "trained" for shooting attacks by playing Call of Duty. Retrieved February 2, 2017, from https://www.theguardian.com/world/2012/apr/19/anders-breivik-call-of-duty

Poels, K., de Kort, Y. A. W., & Ijsselsteijn, W. A. (2007a). *D3.3 : Game Experience Questionnaire: development of a self-report measure to assess the psychological impact of digital games*. Eindhoven. Technische Universiteit Eindhoven. Retrieved from https://pure.tue.nl/ws/files/21666952/Fuga_d3.3.pdf

Poels, K., de Kort, Y. A. W., & Ijsselsteijn, W. A. (2007b). "It is always a lot of fun!": Exploring dimensions of digital game experience using focus group methodology. In *Proceedings of the 2007 conference on Future Play* (pp. 83–89). Toronto: ACM. https://doi.org/10.1145/1328202.1328218

Poels, Y., Annema, J. H., Verstraete, M., Zaman, B., & De Grooff, D. (2012). Are You a Gamer? a Qualitative Study on the Parameters for Categorizing Casual and Hardcore Gamers. *IADIS International Journal on WWW/Internet*, *10*(1), 1–16. Retrieved from http://www.iadisportal.org/ijwi/vol10_1.html

Prynne, M. (2014). Call of Duty linked to four teenage deaths, coroner warns - Telegraph. Retrieved February 2, 2017, from http://www.telegraph.co.uk/technology/video-games/10857984/Call-of-Duty-linked-to-four-teenage-deaths-coroner-warns.html

Przybylski, A. K. (2014). Electronic Gaming and Psychosocial Adjustment. *Pediatrics*, *134*(3), e710–e715. https://doi.org/10.1542/peds.2013-4021

Przybylski, A. K., Ryan, R. M., & Rigby, C. S. (2009). The Motivating Role of Violence in Video Games. *Personality and Social Psychology Bulletin*, *35*(2), 243–259. https://doi.org/10.1177/0146167208327216

Przybylski, A. K., Weinstein, N., & Murayama, K. (2017). Internet gaming disorder: Investigating the clinical relevance of a new phenomenon. *American Journal of Psychiatry*, *174*(3), 230–236. https://doi.org/10.1176/appi.ajp.2016.16020224

Qin, H., Rau, P. P., & Zhong, H. Q. (2007). Construction of online game addiction based on player experience. In M. Helander, M. Xie, R. Jiao, & K. C. Tan (Eds.), *2007 IEEE International Conference on Industrial Engineering and Engineering Management* (pp. 432–436). Singapore: IEEE. https://doi.org/10.1109/IEEM.2007.4419226

Rheinberg, F. (2015). *The Flow-Short-Scale (FSS) translated into various languages*. Potsdam. https://doi.org/10.13140/RG.2.1.4417.2243

Rheinberg, F., Vollmeyer, R., & Engeser, S. (2003). Die Erfassung des Flow-Erlebens. In J. Stiensmeier-Pelster & F. Rheinberg (Eds.), *Diagnostik von Selbstkonzept, Lernmotivation und Selbstregulation (Tests und Trends)* (pp. 261–279). Göttingen: Hogrefe.

Riva, G., Botella, C., Baños, R., Mantovani, F., García-Palacios, A., Quero, S., … Gaggioli, A. (2015). Presence-Inducing Media for Mental Health Applications. In *Immersed in Media* (pp. 283–332). Cham: Springer International Publishing.

https://doi.org/10.1007/978-3-319-10190-3_12

Ross, D. R., Finestone, D. H., & Lavin, G. K. (1982). Space invaders obsession. *JAMA*, *248*(10), 1177. https://doi.org/10.1001/jama.1982.03330100017009

Runic Games. (2012). Torchlight II [Video game]. Runic Games. Retrieved from https://www.runicgames.com/torchlight2/

Runic Games. (2013). Torchlight II GUTS [Development Tool]. Runic Games. Retrieved from https://www.runicgames.com/blog/2013/04/01/guts/

Russoniello, C. V, O'Brien, K., & Parks, J. M. (2009). The effectiveness of casual video games in improving mood and decreasing stress. *Journal of Cyber Therapy and Rehabilitation*, *2*(1), 53–66. Retrieved from https://www.researchgate.net/publication/289131468_The_effectiveness_of_casual_video_games_in_improving_mood_and_decreasing_stress

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, *55*(1), 68–78. https://doi.org/10.1037//0003-066X.55.1.68

Ryan, R. M., Rigby, C. S., & Przybylski, A. K. (2006). The Motivational Pull of Video Games: A Self-Determination Theory Approach. *Motivation and Emotion*, *30*(4), 344–360. https://doi.org/10.1007/s11031-006-9051-8

Salen, K., & Zimmerman, E. (2003). *Rules of Play: Game design fundamentals*. London: MIT Press. Retrieved from https://mitpress.mit.edu/rules

Salen, K., & Zimmerman, E. (2004). *Rules of Play: Game Design Fundamentals*. London: MIT Press.

Sanchez, X., Moss, S. L., Twist, C., & Karageorghis, C. I. (2014). On the role of lyrics in the music–exercise performance relationship. *Psychology of Sport and Exercise*, *15*(1), 132–138. https://doi.org/10.1016/j.psychsport.2013.10.007

Schüler, J. (2012). The dark side of the moon. In S. Engeser (Ed.), *Advances in Flow Research* (pp. 123–137). New York, NY: Springer. https://doi.org/10.1007/978-1-4614-2359-1_7

Senn, S. (1993). *Cross-Over Trials in Clinical Research.* Chichester: Wiley. Retrieved from http://www.jstor.org/stable/2983088?origin=crossref

Seok Hwai, L. (2015). Taiwan revises law to restrict amount of time children spend on electronic devices. Retrieved February 7, 2017, from http://www.straitstimes.com/asia/east-asia/taiwan-revises-law-to-restrict-amount-of-time-children-spend-on-electronic-devices#xtor=CS1-10

Sherry, J. L. (2004). Flow and Media Enjoyment. *Communication Theory*, *14*(4), 328–347. https://doi.org/10.1111/j.1468-2885.2004.tb00318.x

Slater, M. (1999). Measuring Presence: A Response to the Witmer and Singer Presence Questionnaire. *Presence: Teleoperators and Virtual Environments*, *8*(5), 560–565. https://doi.org/10.1162/105474699566477

Slater, M. (2003). A Note on Presence Terminology. *Presence Connect*, *3*(3), 1–5. Retrieved from http://s3.amazonaws.com/publicationslist.org/data/melslater/ref-

201/a note on presence terminology.pdf

Soper, W. B., & Miller, M. J. (1983). Junk-time junkies: An emerging addiction among students. *School Counselor*, *31*(1), 40–43.

Steam. (2017). Torchlight II Global Game Stats. Retrieved May 21, 2017, from http://steamcommunity.com/stats/TorchlightII/achievements

Steamspy. (2017). Torchlight II: All the data and stats. Retrieved February 22, 2017, from https://steamspy.com/app/200710

Suits, B. (1978). *The Grasshopper: Games, Life and Utopia* (1st ed.). Toronto: University of Toronto Press.

Sweetser, P., Johnson, D., & Wyeth, P. (2013). Revisiting the GameFlow Model with Detailed Heuristics. *Journal of Creative Technologies*, (3). Retrieved from https://ctechjournal.aut.ac.nz/paper/revisiting-gameflow-model-detailed-heuristics/

Sweetser, P., & Wyeth, P. (2005). GameFlow: a model for evaluating player enjoyment in games. *Computers in Entertainment (CIE)*, *3*(3), 1–24. Retrieved from http://dl.acm.org/citation.cfm?id=1077253

Takatalo, J., Ihanus, J., Kaistinen, J., Nyman, G., & Häkkinen, J. (2012). Experiencing digital games. In J. Gackenbach (Ed.), *Video Game Play and Consciousness* (pp. 13–40). Hauppauge, NY: Nova Science Publishers. Retrieved from http://www.scopus.com/inward/record.url?eid=2-s2.0-84896568212&partnerID=40&md5=1eecff11f81e7209b488cb4dccad02dd

Tamborini, R., Bowman, N. D., Eden, A., Grizzard, M., & Organ, A. (2010). Defining Media Enjoyment as the Satisfaction of Intrinsic Needs. *Journal of Communication*, *60*(4), 758–777. https://doi.org/10.1111/j.1460-2466.2010.01513.x

Tamborini, R., Grizzard, M., David Bowman, N., Reinecke, L., Lewis, R. J., & Eden, A. (2011). Media enjoyment as need satisfaction: The contribution of hedonic and nonhedonic needs. *Journal of Communication*, *61*(6), 1025–1042. https://doi.org/10.1111/j.1460-2466.2011.01593.x

Thlabela, K., Roodt, J., Paterson, A., & Weir-Smith, G. (2006). Access to computers and internet. In *Mapping ICT Access in South Africa* (1st ed., pp. 21–28). Cape Town: HSRC Press. Retrieved from http://www.hsrcpress.ac.za/product.php?productid=2192&freedownload=1&js=n

Thompson, E. R. (2007). Development and Validation of an Internationally Reliable Short-Form of the Positive and Negative Affect Schedule (PANAS). *Journal of Cross-Cultural Psychology*, *38*(2), 227–242. https://doi.org/10.1177/0022022106297301

THR Staff. (2013). Sandy Hook Shooter Motivated by Violent Video Games, Norway Massacre. Retrieved February 2, 2017, from http://www.hollywoodreporter.com/news/sandy-hook-shooter-motivated-by-422271

Valve Corporation. (2013). Dota 2 [Video Game]. Valve Corporation. Retrieved from www.dota2.com

Vella, K., & Johnson, D. (2012). Flourishing and video games. In *Proceedings of The 8th Australasian Conference on Interactive Entertainment Playing the System* (p. 3). ACM Press. https://doi.org/10.1145/2336727.2336746

Vermeulen, L., Castellar, E. N., Janssen, D., Calvi, L., & Van Looy, J. (2016). Playing under threat. Examining stereotype threat in female game players. *Computers in Human Behavior*, *57*, 377–387. https://doi.org/10.1016/j.chb.2015.12.042

Vorderer, P., Wirth, W., Gouveia, R., Biocca, F., Saari, T., Jäncke, F., … Jäncke, P. (2004). *MEC Spatial Presence Questionnaire (MEC-SPQ) Short Documentation and Instructions for Application*. Retrieved from http://academic.csuohio.edu/kneuendorf/frames/MECFull.pdf

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*, *54*(6), 1063–1070. https://doi.org/10.1037/0022-3514.54.6.1063

Watson, P. (2011). Rule of thumb on magnitude of effect sizes. Retrieved October 30, 2017, from http://imaging.mrc-cbu.cam.ac.uk/statswiki/FAQ/effectSize

Weibel, D., & Wissmath, B. (2011). Immersion in Computer Games: The Role of Spatial Presence and Flow. *International Journal of Computer Games Technology*, *2011*, 1–14. https://doi.org/10.1155/2011/282345

Weibel, D., Wissmath, B., Habegger, S., Steiner, Y., & Groner, R. (2008). Playing online games against computer- vs. human-controlled opponents: Effects on presence, flow, and enjoyment. *Computers in Human Behavior*, *24*(5), 2274–2291. https://doi.org/10.1016/J.CHB.2007.11.002

Wiemeyer, J., Nacke, L., Moser, C., & Mueller, F. "Floyd." (2016). Player Experience. In R. Dörner, S. Göbel, W. Effelsberg, & J. Wiemeyer (Eds.), *Serious Games* (pp. 243–271). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-40612-1_9

Wiseheart, M. (2013). Effect Size Calculator. Retrieved March 13, 2018, from http://www.cognitiveflexibility.org/effectsize/

Witmer, B. G., & Singer, M. J. (1998). Measuring Presence in Virtual Environments: A Presence Questionnaire. *Presence: Teleoperators and Virtual Environments*, *7*(3), 225–240. https://doi.org/10.1162/105474698565686

Wood, R. T. A. (2008). Problems with the Concept of Video Game "Addiction": Some Case Study Examples. *International Journal of Mental Health and Addiction*, *6*(2), 169–178. https://doi.org/10.1007/s11469-007-9118-0

Wouters, P., van Oostendorp, H., Boonekamp, R., & Spek, E. Van Der. (2011). The role of Game Discourse Analysis and curiosity in creating engaging and effective serious games by implementing a back story and foreshadowing. *Interacting with Computers*, *23*(4), 329–336. https://doi.org/10.1016/j.intcom.2011.05.001

Yee, N. (2006). Motivations for Play in Online Games. *CyberPsychology & Behavior*, *9*(6), 772–775. https://doi.org/10.1089/cpb.2006.9.772

Yee, N. (2006). The Psychology of MMORPGs: Emotional Investment, Motivations, Relationship Formation, and Problematic Usage. In R. Schroeder & A.-S. Axelson (Eds.), *Avatars at Work and Play: Collaboration and Interaction in Shared Virtual Environments* (First, pp. 187–207). London: Springer-Verlag. Retrieved from http://www.nickyee.com/daedalus/archives/02_04/Yee_Book_Chapter.pdf

Yuan, F., & Lee, D.-L. (2014). 한중 온라인 게임 중독 방지 시스템 비교 분석 Online game fatigue system comparative analysis between China and Korea. *Journal of Digital Convergence*, *12*(4), 455–460. https://doi.org/10.14400/JDC.2014.12.4.455 한중

Zynga [Video Game]. (2009). FarmVille. Zynga. Retrieved from https://www.zynga.com/games/farmville

8 APPENDICES

Appendix A PRESENCE AND IMMERSION

Our main body argues, briefly, that immersion in digital games is poorly constructed, and is best describe by flow and presence – immersion's constituent parts. Below we present an overview of the presence construct before critically discussing the immersion construct.

A.1 Presence

Presence is a concept that describes the general sense of "being" somewhere other than where one physically exists. Researchers often use presence to study how humans experience Virtual Environments (VEs) such as those in digital games. We present a brief overview of presence, and then review its relevance and application to digital games literature.

Slater (1999) defines the experience of presence by three phenomena:

1. The defining sense of 'being there' in a virtual environment depicted by technology
2. While in a state of presence, humans will respond to virtual events instead of real events
3. After disengaging from the VE, humans have the sense that they have returned from visiting another place

These definitions carry those connotations of physical dislocation of a singular person. Other definitions define presence as non-mediation; that presence occurs when humans fail to acknowledge the mediation occurring between them and the VE, a narrative, or a scene in a painting (Lombard & Ditton, 1997; Lombard & Jones, 2007). This definition does not specifically describe the sense of physical displacement. Thus, the literature identifies a few distinct presence aspects that give the same general sense of the phenomenon, while enabling nuanced discussion of presence's conditions and effects.

Presence literature presents a nebulous history of redefinition instead of refinement (Lombard & Ditton, 1997; Lombard & Jones, 2007, 2015, Slater, 1999, cf. 2003) that permeated UXG research in a different form as immersion (see Section 2.4.4; Brown & Cairns, 2004; Jennett et al., 2008; Takatalo, Ihanus, Kaistinen, Nyman, & Häkkinen, 2012). Presence literature presents a long history of redefinition instead of refinement (Lombard & Ditton, 1997; Lombard & Jones, 2007, 2015, Slater, 1999, cf. 2003). Lombard and Jones (2015) provide a thorough review of the formal presence definitions in an attempt to organise the existing work, and curtail redefinition, misuse, and misappropriations. They identify seven distinct and well-established presence aspects, *spatial presence*, *social presence*, *self-presence*, *engagement*, *realism*, *cultural presence*, and *parapresence*. We briefly review the first five of these aspects in terms of digital games and their utility in digital game research. We exclude discussion of cultural presence and parapresence as their arguments blur the lines between reality and virtuality, and bring into question the fundamental principles of human experience. See Lombard & Jones (2015) for detailed discussion of cultural presence and parapresence.

Spatial presence is the most common aspect mentioned in the literature and pertains to feelings of transportation or dislocation, where players have a sense of being somewhere other than the location that their body inhabits. Certain definitions expand or limit the requirements for spatial presence. Spatial presence is facilitated by many contemporary digital games, particularly those that take place within large and rich virtual environments that provide enough mediated stimuli to reliably overwhelm real world stimuli.

Social presence describes a sense of presence that players get when interacting with other social entities. Humans using videophones or video conferencing technologies may feel present with the humans on the other end of the connection or planet. This experience is common in MMOs where players experience a sense of presence with the other players, represented by avatars, in the game world. Furthermore, players may experience a sense of presence with the artificial entities that populate a game world, so-called Non-Player Characters (NPCs). Kort, Poels, and Ijsselsteijn (2007) argue that digital games are social presence technologies, noting that many online games – which are currently among the most successful – would fail were it not for the social aspect.

Self-presence relates to how the technologies of virtual environments affect a player’s sense of self. For example, a player has a sense of self-presence when they feel that the experiences of their avatar are their own. Some interesting research indicates that some players do not experience self-presence, but rather feel that their avatars are separate social entities (Banks & Bowman, 2016).

Perhaps the most contentious definitions of presence deal with **Engagement**. This aspect describes a deep involvement with some media and its content, including narratives, characters and the medium’s depictions of other worlds. As with most common-word constructs, Engagement is related to other oft-used constructs, such as immersion - as used by (Slater, 1999) - and flow (Csikszentmihalyi, 1990). Lombard and Jones (2015) argue that engagement, as defined by (Jacobson, 2002; Palmer, 1995), is distinct from established and accepted related constructs including presence, flow, attention and others (p. 25)

Realism pertains to the perceived fidelity of observable mediated events, such as audio-visual stimuli and social interactions. Presence occurs when such events in the virtual environment take precedence over events in the real world. Thus, when mediated events feel like real-world event, presence is more likely to occur. Slater (1999, 2003) calls the experience of event replacement *immersion*; Digital games often have stylized representations of real-world or alien events. In these situations, fidelity is not always required provided the events are plausible, and consistent (M. C. Green et al., 2004). When engagement occurs, players perceive realism. Provided the game does not produce digital artefacts (technical anomalies that produce a break in presence), events within the environment become more important to the player and therefore replace attention to real-world events.

Much of the early UXG research committed errors such as careless misuse of domain-specific terminology that inspired Lombard and Jones' (2015) chapter. They allude to presence because they recognize it as a valid experience in digital game, but generally fail to apply it carefully. For example, Sweetser and Wyeth (2005) discuss *transportation* and *immersion* in GameFlow. However, they provide superficial treatment of both concepts. The term *transportation* relates to spatial presence and engagement. Furthermore, they draw upon Brown and Cairns' (2004) intuitive definition of immersion (in games) that incorrectly incorporates presence

There is very little research pertaining to the lasting effects of presence, whereas flow has known lasting impact on well-being, and affect describes both positive and negative emotional states. Furthermore, Torchlight II is a game that likely facilitates some aspects of presence but is not conducive to those most commonly attributed to player experience: spatial and social presence. The game likely facilitates engagement, but much of this can be better described by flow. Thus, we chose to not measure presence as a factor of player experience in this research.

A.2 Immersion

Around the turn of the century, the broader digital game community (players, reviewers and developers,) adopted the term *immersion* to describe game experiences (Cairns et al., 2014). The term was likely sourced from Virtual Reality and Presence discourse, where it has specific meaning; a situation wherein the sensory stimuli from a virtual environment override those of the real world (cf. previous section; Slater, 1999). Furthermore, flow literature sometimes uses the term to describe the essence of the experience: flow consumes one’s awareness and attention. Amongst the digital game community, immersion includes the connotations of Slater’s immersion; to be enveloped or consumed by the game. However, it also means something else (Cairns et al., 2014).

Colloquial use of immersion relies on an intuitive understanding of any particular game’s intended experience. Open-world RPGs such as Skyrim (Bethesda Game Studios, 2011), provide high-fidelity audio and graphics to help players forget that the game world is virtual. Here, immersion is closest to its presence definiton. Visual-novels or point-and-click adventures, such as The Secret of Monkey Island (Games, 1990), forego graphic realism and develop convincing narratives with multi-dimensional characters to which players become attached. Here immersion describes engagement with a narrative. Beat-em-ups and action games, such as Street Fighter V (Capcom, 1992), provide players with consistent and rewarding game mechanics. In this context, immersion describes a game world’s internal consistency or the fidelity of intended experience. The term is loaded with ambiguity and contextual meaning even when carefully considered. (Adams, 2004; Brown & Cairns, 2004; Ermi & Mäyrä, 2005; cf. Sweetser & Wyeth, 2005; Nick Yee, 2006). UXG research took an interest in the concept, using it and investigating it while propogating this ambiguity. For the remainder of this section we use the term *immersion* when discussing the game community's definition, and we use *Slater's immersion* when discussing immersion as per Slater's (1999) definition.

Adams (2004) provides some of the first nuanced thought about immersion as applied to games. They describe three types of immersion: *Tactical* where players become immersed in a performance; *Strategic* where players immerse themselves in a challenge; and *Narrative* where players become immersed in a rich, deep and believable story (Adams, 2004). Ermi and Mäyrä (2007) developed the *SCI model* of immersion, where SCI is an initialism for their three dimensions of immersion, namely Sensory immersion, Challenge-based immersion, and Imaginitive immersion. Sensory immersion is akin to Slater's immersion and describes the experience where sensory stimuli from a game overpower stimuli from the real world: a player's senses are immersed in technology. Challenge-based immersion describes the experience when challenge and ability are balanced. Here *ability* incorporates the physical skills from Adam's tactical immersion and the cognitive skills from strategic immersion (Cairns et al., 2014). It further parallels the component of flow most commonly invoked for digital games: a balance between the skills a player believes they have, and the challenge they perceive. Imaginative immersion describes the experience of losing oneself to a another world: players invest in the game's world and characters, emoting for them and considering their situation. Players practice parasocial interaction with the non-player characters of the game world; they treat computer entities as they would human enities. The SCI model effectively includes Adam's (2004) immersion types despite having no apparent relation to his work.

Brown and Cairns (2004) endeavoured to find an original, distinct, and formal definition for immersion in digital games. Through a small-sample ($N = 7$) grounded-theory investigation, they found that the term is pervasive amongst gamers, and that players generally use it to describe complete occupation, involvement or engagement with a game. Specifically, they identified three sequentially dependant levels of immersion: *Engagement*, which describes any concious effort or energy put towards playing. Engagement is required for *Engrossment,* an experience highly remeniscient of flow. Engrossment is required for *total immersion,* the most extreme form of immersion, which the researchers call presence. (Brown & Cairns, 2004).

The work meets resistance because their findings are not distinct from other established constructs (Ermi & Mäyrä, 2005; Ijsselsteijn et al., 2007). Engagement has existing definitions in presence literature (Jacobson, 2002; Palmer, 1995) and much like the term *immersion,* Brown and Cairns redefine engagement with respect to presence. Engrossment, for the most part, bears descriptive similarities to the components of flow such as "the game becomes the most important [thing]", "a zen-like state", and "everything else is irrelevant" (Brown & Cairns, 2004, p. 1299) which parallels with intense focus, merging of action and awareness, and a loss of reflective self-consciousness. Finally the authors misappropriate presence (Cairns et al., 2014) when they equate it to total immersion (Brown & Cairns, 2004). Sweetser and Wyeth

(2005) included immersion (as used by gamers) within GameFlow. They describe it as a "deep but effortless involvement [with] reduced concern for self and sense of time" (Sweetser & Wyeth, 2005, p. 4). This definition is three components of flow. Their notion of "effortless involvement" bears thematic similarity to a merging of action and awareness - the component that Csikszentmihalyi (1975) speculates is the clearest sign of flow. Whether these three components alone are enough to facilitate flow, or they describe something less – ostensibly immersion – certainly depends on the games, genres, players or contexts in which they are experienced. Nonetheless, the distinction between immersion and flow is unclear (Ijsselsteijn et al., 2007; Sweetser, Johnson, & Wyeth, 2013).

Subsequent attempts to distinguish immersion from the existing and more robust UXG concepts are unconvincing, particularly their dissociation from flow. They present a very brief review reading of flow research, citing only two 20th century sources and dismissing the state as "an extreme experience" (Cairns et al., 2014, p. 7; Jennett et al., 2008). They argue that immersion is a necessary precursor to flow, and that flow occurs only after an actor experiences all nine-flow components simultaneously and with extreme intensity (Cairns et al., 2014; Jennett et al., 2008). However, there is a high level of agreement on the definition of flow: that it is an interplay of the components, not the simultaneous existence of all (Engeser & Schiepe-Tiska, 2012). Contemporary research has no definitive answer as to whether flow is a continuous or discrete phenomenon, but based on the flow research community's research practices, it should be measured as if continuous. Immersion particular definitions of flow (Cairns et al., 2014; Jennett et al., 2008) simply do not agree with the current understanding of flow. On the other hand, Engeser (2012) suggests that researchers should only consider an experience to be flow once an uncertain threshold of intensity is crossed, but thereafter consider it to be continuous. The immersion argument might be compelling if they were willing to engage with flow research, and consider immersion in terms of low-intesity flow states, particularly in light of Engeser's (Engeser, 2012b) suggestions. However, their intention is to isolate immersion and they fail to do so.

Adams' (2004) full descriptions of *tactical immersion* resemble components of a flow state, where action merges with awareness and players become less concerned with their immediate surroundings. He further describes flow-like components in his descriptions of *strategic immersion*: an intense focus and a sense of control through challenge. Similar errors in terminology litter the literature. Sometimes, presence measures are used to assess immersion (Ryan et al., 2006). Other times researchers rediscover existing constructs: Ermi and Mäyrä's (2005) *sensory immersion* is nearly identical to Slater's immersion (Slater, 1999).

Immersion appears to be the choice term to describe a ubiquitous positive experience of gaming. It shares aspects with many existing active-experience models, specifically flow and presence. All attempts to isolate the construct rely on arguments that identify highly specific differences between immersion and each associated construct; however, some arguments are tenuous while others are incorrect. Nonetheless, because psychometrically sound immersion metrics exist, it is clear that the construct describes a unique experience. Whether immersion is a unique phenomena - an interplay of flow, presence and other experiences - or a colloquial word for an established phenomenon, remains to be proven. Notions of cognitive energy and cognitive demand are common in flow literature, and contemporary player expereince research respectively. It is possible that immersion is an isolated construct that has yet to find its identity, but it is also clear that the current operationalization is unconvincing: it is more defined not by what it is, but by what it isn't. We are not sure what it means to be immersed in a game, if it does not describe components of flow, flavours of presence, or both. Thus, we doubt immersion's utility in behavioural research, at least in its current form. Effective use requires a careful operational definition, and confidence that it is more appropriate than established constructs (Brown & Cairns, 2004, p. 2)

Appendix B PARTICIPANT INFORMATION SHEET AND CONSENT FORM

What is this experiment about? We are interested in the ways in which video game players feel while gaming. This is known as *player experience,* and it describes many things including mood, emotional state and level of engagement

What will participation involve? You will be required to fill out some questionnaires, and play the PC game Torchlight II once a week for three weeks.

How long will participation take? Each session will last at least an hour and a half. This includes about half an hour of questionnaires. You are expected to play Torchlight II for at least an hour in each session, but in some cases, you will decide how long to continue playing after that time is up.

You cannot take part in this study without giving us your explicit consent. By signing this consent statement, you are indicating that you understand the nature of the research study, your role in that research, and that you agree to participate in the research. Please ensure that you agree with all the following points before signing:

- I understand that I am participating in gaming research that will be measuring my psychological state as related to gaming.
- I understand that my identity will not be linked with my data, and that all information I provide will remain confidential.
- I understand that I will be provided with an explanation of the research in which I participated and be given the name and telephone number of an individual to contact if I have questions about the research.
- I understand that certain facts about the study might be withheld from me, and the researchers might not, initially, tell me the true or full purpose of the study. However, the complete facts and true purpose of the study will be revealed to me at the completion of the study session.
- I understand that participation in research is not required, is voluntary, and that, at any point during the session, I may refuse to participate further without penalty.
- By signing this form, I am stating that I am 18 years of age or older, and that I understand the above information and consent to participate in this study being conducted at the University of Cape Town.

The word *computer* shall henceforth refer to personal computers, desktop computers, laptops and notebooks and gaming consoles (Xbox, PlayStation, Nintendo etc.).

Participant Age: ____

How do you rate your experience using computers? (circle)

No experience	Some experience	Moderate experience	Lots of experience	Master

How long have you been using a computer? ______________________________

Do you or your household own a computer? ___________________________________

How many? ____

When did you or your household first acquire a computer? ________________

Give a rough estimate of how many hours per week you spend working on a computer (not counting watching series and movies):__________________________

Do you play computer games regularly? _____________________________________

Please list some of the games that you play or have played: ____

Estimate how many hours per week you spend gaming on average: ____________________________

Look at the list of games/game series listed below. Please circle the ones you have played before, and give a rough estimate of how many hours you have spent playing it.

Diablo (I, II, or III) _________________________

Torchlight (I or II) _________________________

Titan Quest ________________________

Dungeon Siege (I, II, or III) ______________________

Today's Date: ________________

Print your First Name: _____________________________

Print your Last Name: ______________________________

Participant Signature: ____

I________________ (researcher), have witnessed that this form was completed prior to the beginning of the experiment, and by the participant identified above.

Researcher's Signature: __________________________________ Date:__________

This study has no intention of subjecting participants to any form of stress, and therefore should not cause in them any form of distress. However, as the questionnaires are of a psychological nature we are aware that worry and anxiety may be experienced by some. If taking part in this study made you feel distressed in any way and you would like to talk to someone about your thoughts please contact us or the University's Student Wellness Centre at <contact details removed for appendix>

Researcher Details: < details removed for appendix>

Appendix C Flow Short Scale

Adapted from (Engeser, 2012a; 2003; Engeser & Rheinberg, 2008)

Think back to when you were playing Torchlight II today. Think about the period of time before you stopped playing. Indicate to what extent you agree with the following statements as they relate to you during this time period.

value	1	2	3	4	5	6	7
meaning	not at all			partly			very much
1. I feel just the right amount of challenge.	○	○	○	○	○	○	○
2. My thoughts/activities run fluidly and smoothly.	○	○	○	○	○	○	○
3. I don’t notice time passing.	○	○	○	○	○	○	○
4. I have no difficulty concentrating.	○	○	○	○	○	○	○
5. My mind is completely clear.	○	○	○	○	○	○	○
6. I am totally absorbed in what I am doing.	○	○	○	○	○	○	○
7. The right thoughts/movements occur of their own accord.	○	○	○	○	○	○	○
8. I know what I have to do each step of the way.	○	○	○	○	○	○	○
9. I feel that I have everything under control.	○	○	○	○	○	○	○
10. I am completely lost in thought.	○	○	○	○	○	○	○
11. Something important to me is at stake here.	○	○	○	○	○	○	○
12. I won’t make any mistake here.	○	○	○	○	○	○	○
13. I am worried about failing.	○	○	○	○	○	○	○

value	1	2	3	4	5	6	7	8	9
meaning	easy						difficult		
14. Compared to all other activities which I partake in, this one is…	○	○	○	○	○	○	○	○	○
meaning	low								high
15. I think that my competence in this area is…	○	○	○	○	○	○	○	○	○
meaning	too low			just right			too high		
16. For me personally, the current demands are…	○	○	○	○	○	○	○	○	○

Appendix D International Positive And Negative Affect Schedule Short Form

Think about how you are feeling at this moment. Please indicate to what extent you are feeling:						
Dimension	Value Item	1 not at all	2 a little	3 quite	4 very	5 extremely
NA	Upset	○	○	○	○	○
NA	Hostile	○	○	○	○	○
PA	Alert	○	○	○	○	○
NA	Ashamed	○	○	○	○	○
PA	Inspired	○	○	○	○	○
NA	Nervous	○	○	○	○	○
PA	Determined	○	○	○	○	○
PA	Attentive	○	○	○	○	○
NA	Afraid	○	○	○	○	○
PA	Active	○	○	○	○	○

Appendix E BETWEEN-SUBJECTS PILOT STUDY

E.1 Introduction

Problematic gaming is an issue that is gaining attention in both popular media (Seok Hwai, 2015; Yuan & Lee, 2014) and academic literature (Entertainment Software Association, 2016). The problematic behaviour is sometimes labelled as gaming addiction and describes the condition where a person's well-being is negatively affected by the time they spend playing digital games. Many parties involved take the label of addiction literally and intervene using established treatment methods, such as online support forums, Gaming addiction clinics, and various forms of behaviour therapy (Heo et al., 2014). At least two policies designed to influence screen time have been employed in countries including South Korea, China and Taiwan (Yuan & Lee, 2014). Literature exploring the effects that these approaches have on players is scarce, or non-existent. Treatment methods remain a voluntary undertaking, while the regulatory systems of Chinese and South Korean governments are compulsory for all game players and thus the effects are imposed. This pilot study serves as a pilot for a larger study that attempts to identify effects that these two systems have on player experience. Our aim here is to identify unforeseen methodological and practical issues with our system and experiment designs.

Behaviour policies

It is common for various authorities to restrict screen time. Nearly 80% of US parents place time limits on digital game play (Pardo, 2010). In 2012, the South Korean Ministry of Education, Science, and Technology announced a law that limits how much time gamers could spend playing in any one 24-hour period: The game forces players to take a ten-minute break after two hours of play after which they may continue for an additional hour. (Engeser & Schiepe-Tiska, 2012; Russoniello et al., 2009; Vella & Johnson, 2012).

Design policies

The General Administration of Press and Publication (GAPP) in China introduced a policy with the same intention as South Korea's – to limit time spent in game. Much of the available information about this policy is based on news media correspondence. The GAPP recommend that 3 hours is the upper limit of healthy play time, and that after 5 hours, playtime is unhealthy (Kowert & Quandt, 2016). However, instead of policing behaviour by strictly controlling the daily play time, the GAPP insist that developers modify their games design, suggesting that developers place a dynamic handicap upon progress (Thompson, 2007).

Player experience

Player experience is a term used to describe the subjective experience had by an individual because of the virtual and real-world environments. Most interpretations of the construct are recognized to be multi-dimensional, incorporating elements of *immersion*, *presence*, *flow* and *affect*. Flow also deals with engagement, but less with the virtual environment, and more with the percentage of cognitive process that are engaged by the game system. Affect is a psychological term used to describe the subjective experience of emotion, or feelings. Affect is commonly categorized into *positive-affect* (PA), which incorporates such phenomena such as feeling interested, excited, strong enthusiastic et al. and *negative-affect* (NA) which includes phenomena such as feeling distressed, ashamed, jittery and afraid et al. Various reliable user-experience measures exist for a variety of computer systems, but none which are appropriate for the unique, multi-dimensional experience presented by digital games. One measure, the GEQ has attempted to capture all these dimensions in a 36-item questionnaire, but unfortunately there has been no evident validity and reliability testing performed.

We intend measuring those aspects of experience that have a lasting effect on players, namely affect and flow. Research show that both PA and flow are associated with positive well-being (Engeser & Schiepe-Tiska, 2012). Presence generally describes a quality of an experience as it occurs. *Immersion in games* is a new and problematic construct that has yet to distinguish from presence and flow (see Appendix A). They are thus excluded from this experiment because the concern is with the experience of the player immediately before and after the intervention (the instantiation of the conditions).

Research questions

Both types of policies impose functional changes upon players that are not intended by the game designers. As a result, the systems potentially change the way in which players experience their hobby. We are interested if they do, and how they do.

E.2 Methods

Design

We ultimately intend to directly compare each policy's experiential effects. Surveys may provide strong ecological validity, but do not offer the rigorous control over the genre, game, player, and social context factors that are important to player research (Senn, 1993). Our approach in this pilot study uses exploratory user testing and our aim is twofold: to identify potential issues with the systems themselves and to identify extraneous factors that might discourage using laboratory methods. To this end, we employed a single-blind between-subjects experiment in form. All participants played one condition only and were not told that this session would play out differently from any other session, however we did not run inferential statistic tests. Policy type serves as our independent variable whereas affect and flow form our dependant variables. We used the same game to control for game and genre differences. To help control for individual player differences, we measured trait affect for each participant before their session, and state affect after their session. Trait affect is the extent to which participants generally experience the emotions that make up the scale. State affect is the extent to which participants experience the same emotions at a certain point in time.

Experimental system

We used the game *Torchlight II* to test the two conditions. The developers, *Runic Games* provide a content creation toolset *GUTS* that allow people to modify the game in a variety of ways.

We developed a *shutdown* system based on the concept of the behaviour policies (see Section 3.3.1) that strictly controls screen time and allows users to play for a certain amount of time, after which the game shuts down, and players are not allowed to interact with the game at all. We implemented this system with a simple Windows batch file. Running the batch script file opens the game, waits for an hour before killing the Torchlight2.exe process.

```
1   "<PATH_to_Torchlight_II_Directory>\Torchlight2.exe"

2   ping 127.0.0.1 -n 3700

3   c:\windows\system32\taskkill /im Torchlight2.exe

4   c:\windows\system32\taskkill /im ModLauncher.exe
```

Figure 8.1 Batch script for shutdown system. Line 1 calls the Torchlight II executable. Line 2 pings local for 3700 seconds before lines 3 and 4 can run, which then kill the tasks associated with Torchlight II and effectively shut the session down.

We developed a *fatigue* system based on the design policies reported in 8.1. Our system allows users to play Torchlight II unimpeded for an hour, after which the reward structure of the games begin to fatigue, reducing the drop rates of gold and magical items. Additionally, the amount of experience gained for each enemy slain becomes negligible. This system allows players to continue playing the game, but character progression is stunted.

We modified Torchlight II to include our fatigue system using the game's content creation tool *GUTS.* After one hour the game would fatigue slightly with every instance of damage that an enemy received. The Gold drop-rates, magical-item drop-rates and experience gain-rates were decreased by 0.7%. For our first participant, the various attributes that define a character's overall ability to defeat enemies (Strength, Dexterity, Focus and Vitality), and the speed at which the player could move was decreased by 0.3 units. After the first session using this intervention, the implementation was changed. The attribute and speed fatigue were removed because they removed the player's ability to continue playing. See Appendix E for more discussion.

Measures

We used the International Positive and Negative Affect Schedule Short Form (Thompson, 2007) to measure affect. The instrument uses Likert-type scales to measure the extent to which respondents feel positive: *alert*, *inspired*, *determined*, *attentive* and *active*. It further measures the extent to which they feel negative: *upset*, *hostile*, *ashamed*, *nervous* and *afraid*.

We used The Flow Short Scale (FSS) (Rheinberg, Vollmeyer & Engeser, 2003; cf. Engeser & Rheinberg, 2008) to measure flow, as this scale has been shown to be reliable and valid, and is useful for measuring the intensity of flow in each activity.

We transcribed the I-PANAS-SF and FSS to the online survey tool *LimeSurvey* where presentation and data capture occurred. The tool and database are hosted on the experimenter's university server, and can be accessed from any location with an internet access point.

Participants

At the time of this pilot study, the intention was to use just two conditions for the experiment: The fatigue group, and the shutdown group. Four participants where used, split evenly between the two conditions. Participants are male aged between 23 and 26. These participants were selected for convenience, rather than specifically for appropriateness. None the less, all participants represent some portion of the gaming population; they are experienced gamers who at the time of the study play games for at least five hours per week, and for more than four hours in one sitting at least once a month.

Procedure

All sessions were run in the participant's own home. All were quiet rooms far from distractions such as human traffic and television. All four sessions had the same basic structure. Participants were asked to navigate to the website hosting the pre-test I-PANAS-SF, and complete the schedule. The pre-test measures the participant's affect in general. Upon completion, they were presented with instructions to begin the game. For the fatigue condition, the modified game was already running as a background process and merely required that the participant click the Torchlight2 icon on the Windows 7 taskbar, and then click the button labelled "play with mods". For the shutdown condition, the participants were asked to double-click the Torchlight2.bat file, and then press the "play button".

All participants were then instructed to start a new game, and create a new character; a mandatory requirement for beginning a new game. Participants were not told which condition they would play. Each participant then played for the duration of the session, either until the game shutdown, or became too fatigued for them to continue as was the case with participant FH (see Appendix E). Participants were then asked to navigate to the post-test I-PANAS-SF website and complete the schedule. The post-test measures the participant's affect at that specific point in time. When the post-test is complete, LimeSurvey automatically navigates to the page hosting the FSS, and participants are asked to complete that in context of the game they have just finished playing.

E.3 Qualitative Results

Session 1 - Participant: FH. Condition: Fatigue

Participant FH is a veteran Torchlight II player, having spent 98 hours playing at the time of the experiment.

This first session was handled like a single blind experiment: FH was not informed that a modification was in place, that there were fatigue elements in the modification, nor how they worked. FH was asked to “play until you want to stop.”

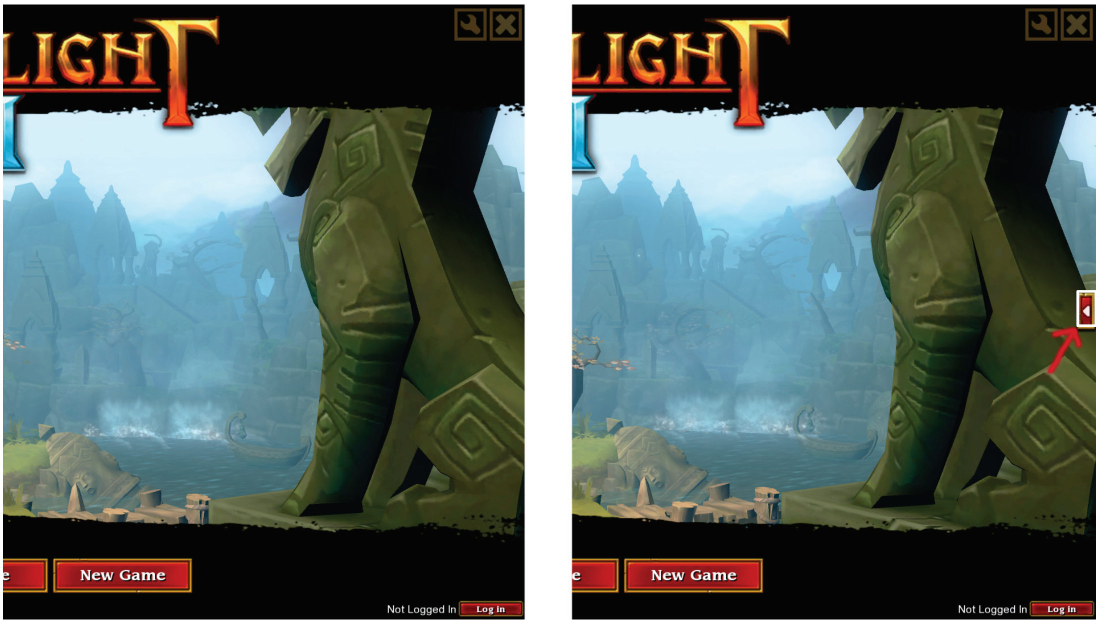

Figure 8.2 Differences on Torchlight II start screen for fatigue and control conditions.The left image does not have a mod installed, whereas the right-hand image does. The “Mod Details” button is visible in the right-hand image on the right side of the screen

FH noticed the “Mod Details” (see Figure 8.2) button on the start screen that is not there unless a modification is installed and was thus actively looking to see what had been changed. Comments made during the session showed that they were expecting very noticeable changes to the game.

“That pile of bones is new. Did you put that there?”

FH also had limited knowledge of the overall intentions of this project, and projected a difficulty increase onto the game when there was none. This comment was made after about 5 minutes of gameplay, long before the intervention begins to have an effect the game:

“I can tell you have made this significantly harder. I should be one-hit killing these Ratlins (enemy type) but it is taking two or three hits.”

After the hour, the fatigue elements started being applied. The most notable element was the decreasing of movement speed. At first, FH, assumed that their character was under the effect of a slowing spell, and that it would wear off. Within about two minutes of the intervention, FH realized the slowing effect was not going away and stated

"I can't play anymore. This is unplayable."

Thus, the session finished.

Discussion

The walking speed fatigue was the most prominent aspect of the system presented to FH. FH was never aware of the fatigue placed upon the gold and item drop-rates, and did not notice that their character had ceased to gain experience. Most surprisingly, they did not notice that their character's attribute statistics were falling even though the numbers are displayed numerically to players and they were using those numbers to choose which items to equip. The walking speed fatigue clearly dominated their attention.

While this session was intended to be a fatigue condition, it was effectively a shutdown condition; FH's ability to continue playing was removed. At first, it was thought that by editing the rate at which movement speed was fatigued, this problem would be removed. But the intention of the fatigue system is not to remove the player's ability to continue playing, but rather to encourage the player to take a break. Thus, all elements of fatigue not related to reward (movement speed and attribute fatigue) were removed from this condition. Furthermore, we decided to remove the single-blind aspect of the sessions because FH spent so much time and effort looking for the modification. Had they known what would happen, then they may have focused more on playing the game. The real-life policies are public knowledge, and those players are aware of how the systems function. This decision would increase the ecological validity of further sessions.

Finally, because of the comments being made, we decided that all comments should be formally recorded, and that the time and length of the session should be recorded to correlate with comments, and to measure how long after the intervention instantiates the player chooses to stop.

Session 2 -Participant: RP. Condition: Fatigue (adapted)

Participant RP had never played Torchlight II before, but had spent an estimated 140 hours playing games of the same genre in the past year alone; Action RPG/Hack-and-Slash RPG. Thus, they were familiar with the game's mechanics.

This was not run as a single blind study. RP was informed that after an hour of play, the rate at which they acquired experience, gold and magical items would decrease, but that they could continue playing for as long as they cared to. RP had no in-game indication of time progression, nor the point at which the intervention was instantiated. Thus, the session was timed manually by the experimenter. RP played for an hour and thirty-one minutes.

After forty-seven minutes of gameplay, RP indicated that he thought the intervention had begun by saying

"Ah. I'm not getting much experience anymore…"

Interestingly, this did not seem to affect his desire to play as he continued playing for an additional forty-four minutes. After this time, he stated

"Ok, I'm done."

Other than these, RP did not make many comments specific to the system, or any that were directed at the experimenter. Most comments were para-social interactions; directed at the game's characters and enemies.

Discussion

The early assumption of intervention means players might try to stop playing before the fatigue system is instantiated. This implies a potential placebo effect associated with the knowledge that the system will fatigue. Furthermore, participant RP did not stop when they believed their character progression had stopped. In fact, they nearly doubled their session's length. It remains unclear as to why they chose to stop at this point. Because of this, we shall record a player's reason for hiatus for every session to find out why players are motivated to stop playing, especially if the fatigue system is not the cause for their hiatus.

Session 3 – Participant: DP. Condition: Shut down

Participant DP, much like RP, had not played Torchlight II before, but had played similar games and was familiar with the game mechanics. They were informed that they had exactly an hour to play from the time they started the game after which the game would exit. There was no in game indication of how time was passing.

DP played very quietly, with no comments made during the game session, but a single exclamation when the time was up.

"Aw, damn."

When asked what was wrong, DP responded

"I didn't get to pick up the rare item on the ground."

Discussion

Torchlight II saves the game state upon exit (even when the batch command *taskkill* is invoked, as is the case with this implementation). When a game is loaded, it loads the exact state that it was left in. Most games of a similar genre, particularly MMORPGs, save the character state (stats, skills and item in inventory), but not the game state (items on the ground/enemies slain). In the case above, if DP returned to the game, they would have been able to retrieve the rare item that was on the ground. In other games, that item would have been lost, and frustration may then result when the player has invested much time in the game.

Session 4 – Participant: RM. Condition: Shutdown

Participant RM had spent 5 hours playing Torchlight II, and had prior experience with the genre. They were informed that they had exactly an hour to play from the time they started the game after which the game would exit. There was no in game indication of how time was passing, but RM asked if they could look at their watch while playing to monitor the time. It was allowed.

After 50 minutes of playing, RM announced,

"Time to prepare for the end."

And then returned his character to that act's town. The town is a place in the game world where there are no enemies. In towns, there are shops where players can buy and sell items, and there are chests were a player can store items that they have found/bought for later use, or to make room in their own inventory. RM spent the next five minutes organizing their inventory and chests. When finished, they just ran around and browsed shops until the game exited. When questioned about this behaviour, RM said,

"I like things to be in order before I stop playing, otherwise it doesn't feel like a real…um... ending point. I don't like to leave my character in some sort of limbo until I return. If I could have put my character to bed, like in Skyrim, that would have been best."

Discussion

The desire to have everything in order before exiting the game is very interesting. It is as if RM needs to create a specific hiatus point in their gameplay session, or else they would feel uncomfortable, and anxious. This only furthers this experimenters desire to explore "normal" hiatus behaviour. To this end, we shall ask participants where they were in the game world when their session came to end, and what it was that they were doing. If players tend to end their sessions in some sort of safe space where they are prepared for their next session, such as in town or in bed as per RM's comments.

E.4 Quantitative results

Here we present graphs describing the results of our quantitative measures.

Affect

We used the *International Positive And Negative-affect Schedule Short Form* (I-PANAS-SF) to measure trait and state affect in our participants.

Figure 8.3 (p. 122) tells two stories. Participants FH and DP show increased positive-affect after their session came to end, although DP's score increase by one point only. This suggests that playing Torchlight II may help these participants experience positive emotions. Participants RP and RM show decreased positive-affect, suggesting that the game and our system impedes the experience of positive emotions.

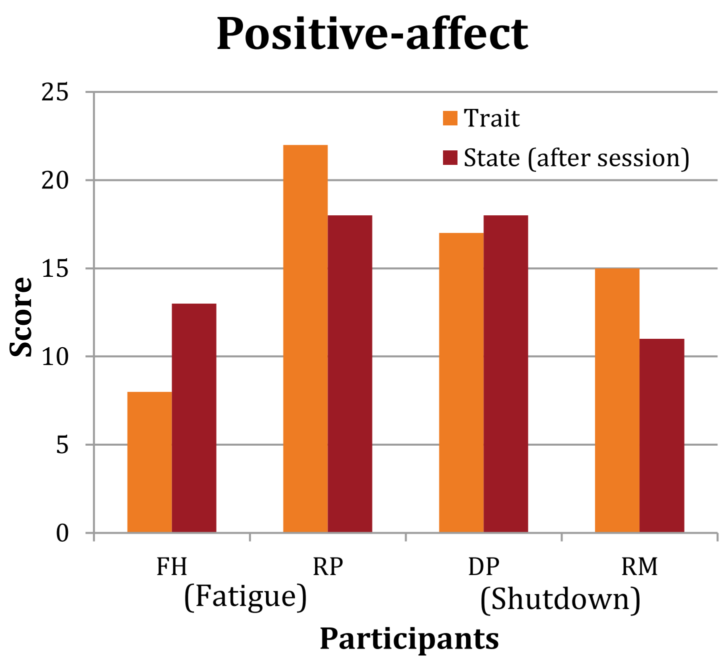

Figure 8.3 Comparison of trait positive-affect (orange) and state positive-affect (red)Score range is 5-25; less is better.

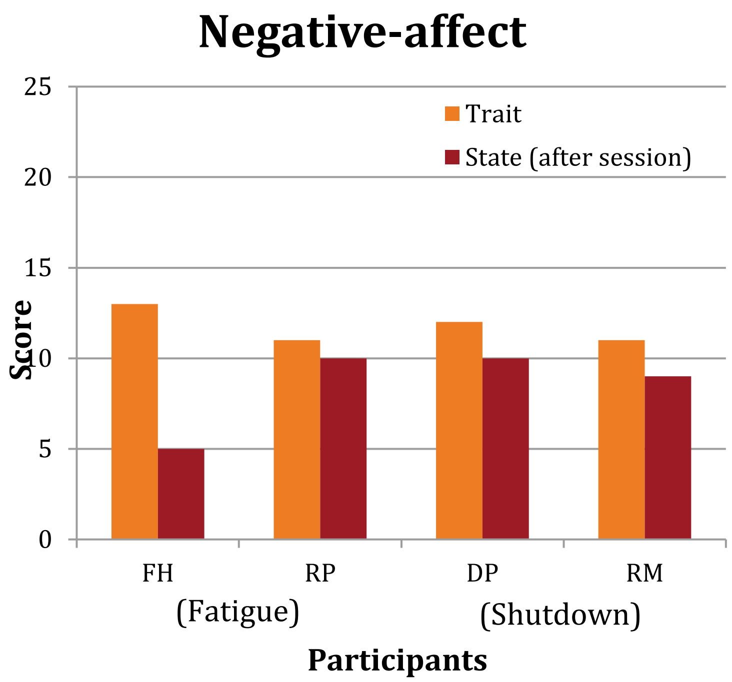

Figure 8.4 Comparison of trait negative-affect (orange) and post-game state negative-affect (red).Score range is 5-25; less is better.

Figure 8.4 shows that our participants tend to experience fewer or weaker negative emotions after using our systems than they tend to experience in everyday life. FH shows the most drastic decrease, whereas the rest change by two points on the scale at most. This suggests that both systems allow players to offset negative emotions.

Discussion

Only FH and DP present the best-case scenario for affect; an increase in PA and a decrease in NA. Their data suggests that the systems increase positive emotional experience while offsetting negative emotional experiences. The data generated by RP and RM does not conform to this ideal; their data suggests the systems reduced PA for these participants. However, we are unable to conclude so. Our PA results showed vastly different trait scores and state scores highlighting the role that individual personal differences play in subjective measures. Our intention was to normalize each participant's scores by taking the difference between trait and state scores, and then compare the normalized scores from the shutdown and fatigue systems. These results demonstrate a methodological weakness with this approach. Although RP's state score was lower than their trait scores, it was still high. The normalized score suggests that RP's session decreased positive-affect from their normal experience, which is ostensibly a bad thing, but their positive emotional experience is still strong. Thus, this approach misrepresents the experience of players by categorizing a positive experience as negative. Due to these discrepancies in scores, this normalization offers insufficient control over individual personal differences. We must revise our experiment design.

Flow

We used the *Flow Short Scale* (FSS) to measure the intensity of flow experience (see Figure 8.5 p. 124) and some of its components that are important to gaming as an activity namely the *perceived importance* of the activity to the actor (see Figure 8.6 p. 124), and the *demand*, *skill*, and *demand-skill balance* (see Figure 8.7 p. 125). Unlike affect, flow is a state only, and not a trait. Therefore, we only measured it after each the game session.

Flow intensity (see Figure 8.5 124) shows to what extent participants experienced a flow state. Flow scores are relatively high, with *absorption by activity* contributing the most, relatively, in all but RP's session. It is important to take note of flow's two dimensions fluency and absorption, to understand which factor is contributing most to the participant's scores. While we might not use this data, it might be useful in explaining anomalous data.

Perceived importance is the value that players associate with success and failure in the activity. Scores here are relatively low, with RP scoring the lowest. Their result is unsurprising given their low absorption scores (see Figure 8.5). There is no clear difference between the conditions, although the shutdown condition does boast higher scores.

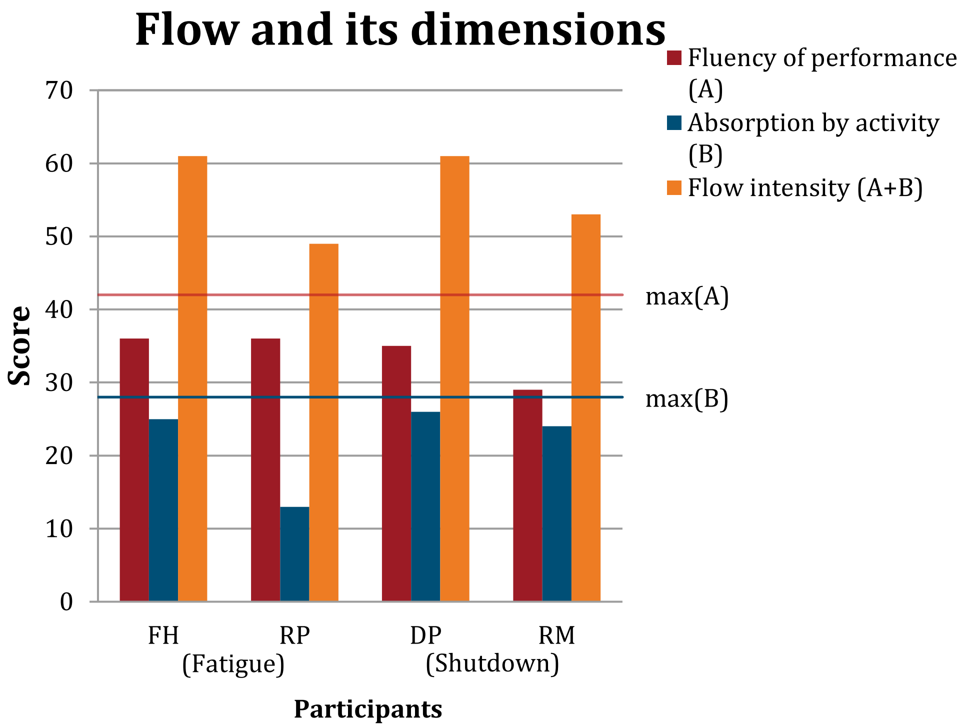

Figure 8.5 Flow results for each participant.Flow intensity (orange; range 10-70) comprises of Performance (red; range 7-42) and Absorption (blue; range 4-28). Red and blue lines show max values for respective dimensions. More is better

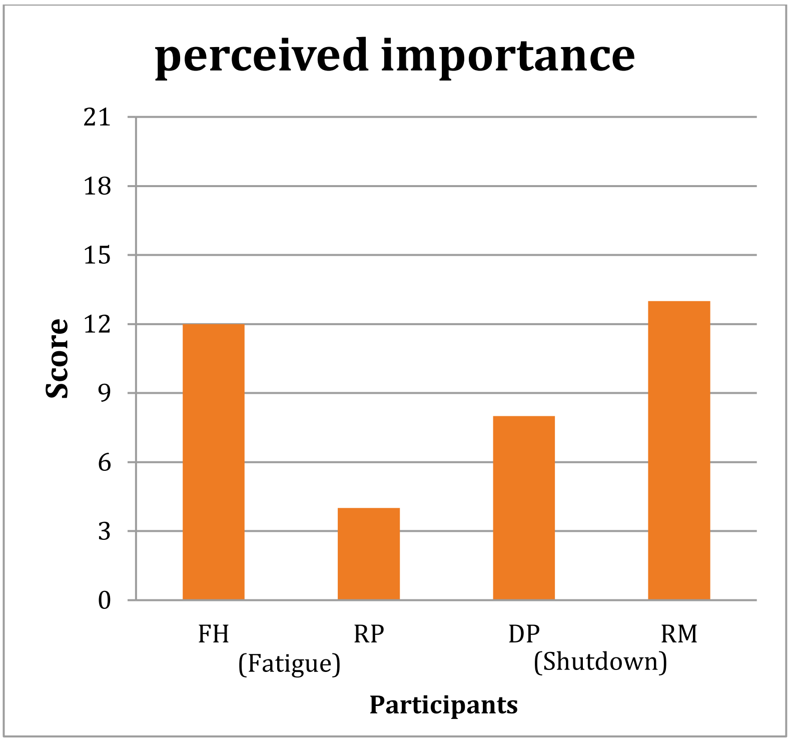

Figure 8.6 Perceived importance scores.Graph displaying how important a successful outcome of the game was to each player. Score range 3-21.

This graph shows an interesting pattern for FH and RP. Both feel that the game under the fatigue condition demands barely any skill from them, and that their personal skill at the game is above average (score of 5) and very high. And yet, they feel that this is in balance. DP also feels that the game's demands are just right for their skill level, but they indicate that the game's demand is on the difficult side. Of the four participants, FH and RP had the most experience in the genre. It is possible that the novel aspects of the game were demanding for DP as they needed to learn and execute new actions.

Discussion

Flow is a positive experience strongly associated with positive well-being (Girden, 1992). We therefore expect that high flow scores will correlate high PA, and low NA. It might be interesting to see if this assumption holds true. With the data that can be gathered from this experiment we can do so.

Our flow results confirm our suspicions regarding our experiment design. The variation in scores for the FSS show that flow experience is not uniform, and does change with individual player differences. This is especially apparent in the demand-skill results (see Figure 8.7). If we continue with our between-subjects design, we will not be able to draw useful conclusion when we compare conditions.

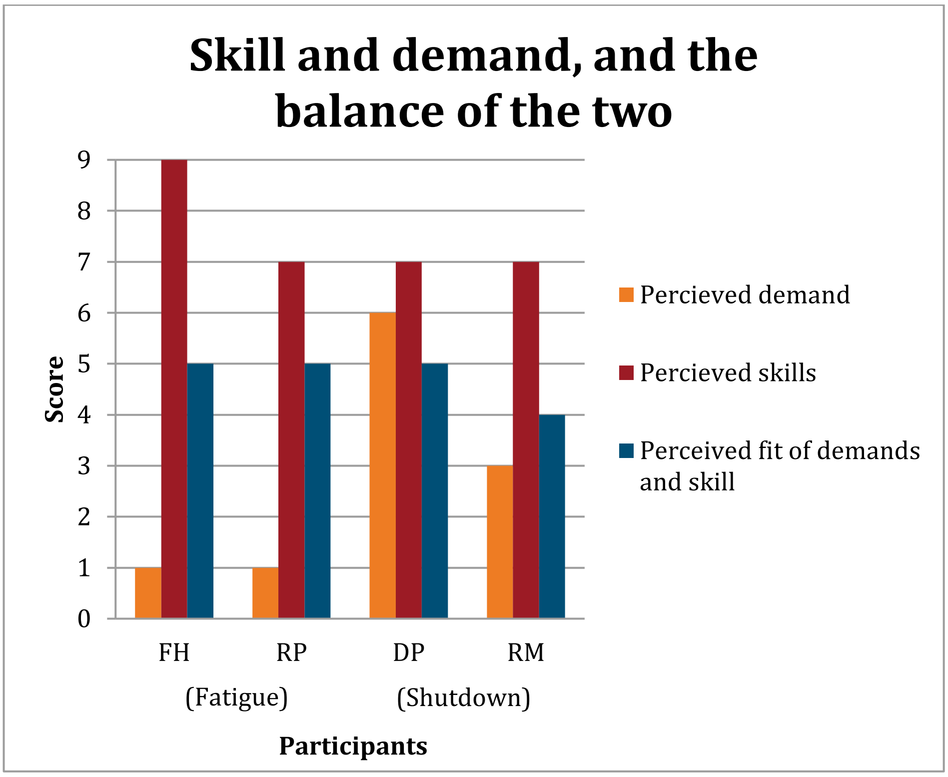

Figure 8.7 Frequency distribution for demand, skill, and perceived fit of the two.Perceived demand (orange) shows how much skill the participants feel that the game demands of them relative to other activities (1 = "easy"; 9 = "difficult"). Perceived skill (red) reflects how much skill they feel they have at their disposal (1= "low"; 9 = "high"). Perceived fit (green) shows how high the demands of the game are for the players (1 = "too low"; 5 = "just right"; 9 = "too high").

E.5 Conclusion

We need a new experiment design to better represent the experiences of our systems to make meaningful comparisons between them; to accommodate the additional conditions and dependant variables, and to better control for individual personal differences. Thus, our next experiment will use a repeated measures design, and borrow design elements from a cross-over design (Senn, 2012) commonly used clinical trials where participants are exposed to every condition and they each serve as their own control. In our between-subjects pilot (Appendix E), we identified the need to record all comments that the participants make during a session, measure additional factors, including session length, and reason for hiatus. Appendix H contains the details of the pilot study that incorporates these changes.

Appendix F ETHICS CLEARANCE

Faculty of Science
University of Cape Town
RONDEBOSCH 7701
South Africa

E-mail: richard.hill@uct.ac.za
Telephone: + 27 21 650 2786
Fax: + 27 21 650 3456

29 April 2014

Mr Bryan Davies
Department of Computer Science
University of Cape Town

Dear Mr Davies

The User Experience of Hiatus Events in Video Games

I am pleased to inform you that the Faculty of Science Research Ethics Committee has approved the above-named application for research ethics clearance, subject to the conditions listed below. You are required to:

- implement the measures described in your application and in the further information submitted to the committee which outlines changes to the research design, to ensure that the process of your research is ethically sound

- uphold ethical principles throughout all stages of the research, responding appropriately to unanticipated issues: please contact me if you need advice on ethical issues that arise.

Your approval code is: FSREC 013 – 2014

I wish you success in your research.

Yours sincerely

Dr Richard C Hill
Chair: Faculty of Science Research Ethics Committee

Appendix G Additional Questions

For each numbered item, we requested a short description. For each bulleted item, participants were asked to select one.

1. Please tell us why you stopped playing today

2. Where were you when you stopped playing? In the space provided, please briefly explain what sort of activities you doing in that area.
   Choose one of the following answers
   - In a town/camp
   - In the wilderness
   - In a dungeon

3. How badly do you want to continue playing? Please explain your choice in the space provided.
   Choose one of the following answers
   - I never want to play again!
   - Maybe sometime in the future
   - Definitely sometime this week
   - Later this week
   - Later today

I want to continue playing immediately!

Appendix H CROSS-OVER PILOT

H.1 Introduction and aims

Our research intends to investigate and compare the user-experiences associated with session length policies such as those practiced in China, South Korea and Taiwan. The behaviour policy strictly regulates session length and shuts down games when a player's time is up. The design policy has fatigue mechanics built into the game that modify the rate at which players gain rewards or build their characters. The fatigue mechanics usually begin affecting the game after a period.

Our previous pilot experiment used a between-subjects design. The results and lessons learned in that study made it clear that our design was flawed. It did not allow us to make meaningful comparisons between the policies due to individual personal differences that affect the measured outcomes of subjective user experience. The design further lacked a condition which could represent a normal gameplay session. The pilot study also lacked a measure of some important aspects of player experience including participant comments, session length, and reasons for stopping. This appendix presents a revised experiment design and the data from the associated pilot study. From this pilot study, we refined our design into what is the main experiment (see Section 3).

Our aims for this pilot study are two-fold. We wish to use it a test-run for the final experiment. Using these results, we will review and refine our design, or rethink our data collection methods. We also wish to collect data to calibrate two aspects of our final experiment, should we continue with it: The length of unpoliced play time granted to players before their session either shuts down or begins to fatigue. We also wish to estimate the sample size that will give us adequate statistical power.

H.2 Design

We adopted a repeated measures experiment design using methods from a cross-over study based on the results and lessons learned from our previous pilot study (see Appendix E). Any player's user-experience is influenced by the tropes of the *genre* they are playing, and the unique qualities of the specific *game* they are playing. The experience is further influence by individual *player* characteristics, and the dynamics of the social and physical *context* in which the player plays a game. An experiment allows us to control the *game* and *genre* variables by making all participants play the same game in a similar way. We can control many player and context variables by screening participants, maintaining a consistent schedule and environment, and controlling the social context in which participants play. The player factors are further controlled through the features of a *cross-over* design (Thompson, 2007) where the effects of participant variables on the measured outcome are minimized with each individual serving as their own control (Cohen, 1988). Participants experience each condition at least once to represent each exposure. The design must vary the treatment sequence and must include 'washout' period between each treatment to minimize carry-over effects and interaction effects. This is important for our research because of the potential influence of player factors upon the effects identified in our previous pilot (see Appendix E for details of results and discussion).

Independent variables

Policy acts as our independent variable and has three levels, or experimental conditions: *shutdown*, *fatigue,* and *control*. The *shutdown* condition simulates the essence of the behaviour policies described in Appendix H.1 by forcing the end of any given session after a specified period. The *fatigue* condition represents the design policies as described in Appendix H.1. Players experience unmodified gameplay for a period, after which character progression rates and loot-drop rates tend towards zero. The remaining level, *control,* represents the absence of a hiatus policy: other than those intended by the game designers, the condition does not manipulate game mechanics, does not impose a session length upon players. This condition is sufficiently different from the shutdown and fatigue conditions to act as a baseline against which we can compare results.

Dependant variables

We measure *session length*, dimensions of *player experience* (flow and affect; see Section 2.4 for detailed discussion of player experience). Williams (as cited in Cochran & Cox, 1957) argue that carry-over effects are inherent in cross-over studies and are difficult to control for. The same goes for individual personal differences of participants. We therefore also ask participants to comment on their decision to end their session, we record conversations and observable behaviours during each session that will explain or allow commentary on any outlying behaviour or data.

Hypotheses

Our experiment aims to test the following hypotheses:

a. The type of policy has an effect on Positive-affect.
b. The type of policy has an effect on Negative-affect.
c. The type of policy has an effect on the experience of flow state.
d. The type of policy has an effect on game sessions-length.

H.3 Materials

Experimental system

See Appendix E for details. Our experimental system remains largely unchanged from the previous study and did not change for our main experiment. For the shutdown condition we ran the Torchlight II process through a Microsoft Batch script that counted out 3,700 seconds before killing the process (see Figure 8.1 p. 116). For the fatigue condition we modified Torchlight II using the content creation tool, GUTS[35].

Table 8.1 Values governing the fatiguing of loot. Values chosen so that loot rates reach (effectively) zero at between 63 and 65 minutes.

Fatigued stat	Effect Name (as per GUTS)	Amount
XP Gain	XP GAIN BONUS	-5%
Magical drop	PERCENT MAGICAL DROP	-20%
Gold found	PERCENT GOLD DROP	-10%
Gold drop	GOLD DROP	-10%

Our modifications allowed players a given period of unhindered gameplay before fatiguing the game's loot system. After this period, the rate at which players gain loot decreases until the loot systems effectively halt. At the beginning of a game session in which the modification is active, a timer is set to 3700 seconds - a little over an hour - to account for time spent loading the game and navigating the menu. After the timer runs out, the variables governing the loot probabilities decrease by a percentage (see Table 8.1) with every instance of damage dealt by the player's character to an enemy. Through this, the variables very quickly tend towards zero, severely stunting the rate at which players acquire fresh loot. These fatigue rates differ from those in the original pilot study, and make it so that drop rates become effectively non-existent within three minutes of gameplay.

The control condition required no modification. Players merely played the unmodified version of Torchlight II.

Measures

We used the International Positive and Negative-affect Schedule Short Form (I-PANAS-SF) (Thompson, 2007) to measure affect. The instrument uses Likert-type scales to measure the extent to which respondents feel positive: *alert*, *inspired*, *determined*, *attentive* and *active*. It further measures the extent to which they feel negative: *upset*, *hostile*, *ashamed*, *nervous* and *afraid*.

We used The Flow Short Scale (FSS) (Rheinberg, Vollmeyer & Engeser, 2003; cf. Engeser & Rheinberg, 2008) to measure flow, as this scale has been shown to be reliable and valid, and is useful for measuring the intensity of flow in each activity.

We used a clock to measure session length and rounded up to the nearest minute. Higher levels of accuracy for this session length are inconsequential.

We transcribed the I-PANAS-SF, FSS, and all additional items to the online survey tool *LimeSurvey* where presentation and data capture occurred. The tool and database are hosted on the experimenter's university server, and can be accessed from any location with an internet access point.

[35] Runic Games, 2013

Table 8.2 Policy-exposure order. Show the three participant groups, and the sequence in which they experienced each policy.

Participant group	First session	Second session	Third session
Sequence SCF	Shutdown	Control	Fatigue
Sequence FSC	Fatigue	Shutdown	Control
Sequence CFS	Control	Fatigue	Shutdown

Participants

We recruited from the University of Cape Town ten male computer science students aged between 22 and 27 years who played games an average of $M = 12.4$ hours per week. All had experience playing Torchlight II or games that fall into the same genre having spent an average of $M = 103.6$ hours in game. Participants were required to attend three play sessions and were randomly assigned to one of three balanced participant groups. Each group experienced the three conditions in a unique sequence. This partially counterbalances the order of exposure to help minimize and identify order effects (see Table 8.2). Participant groups SCF and FSC both had three members, while group CFS had four.

Procedure

For the first session only, before it began in earnest, each participant read through a document informing them of their role, requirements, anonymity, rights and recourse for the experiment. If they gave us their consent to continue, they were asked for some demographic information including relevant gaming history (see Appendix B). For all sessions, participants were informed about the policy they would experience as follows:

- Shutdown condition: You are allowed to play for one hour. After this time the game will save your progress and exit. You will not be allowed to continue.
- Fatigue condition: You are allowed to play for as long as you like. After an hour, the rate at which you get experience, gold, and magic items will slow down. You may continue to play even after this happens.
- Control condition: You must play for at least an hour but as long as you like.

Once the session had come to an end, the experimenter recorded their session length, and the participants were directed to an open web-browser containing the LimeSurvey website. Here, they completed the affect schedule and flow scale, and described why they chose to stop playing where appropriate.

After the final session, participants were thanked for their help, and given the opportunity to ask any questions, or voice any concerns that they might have had. If any participants had felt distressed by the experiment for any reason at any stage and wished to desist, they would have been allowed to and advised to seek assistance from the university's student health services or would have been referred to a clinical psychologist.

H.4 Results

We downloaded all participant data from our survey hosting LimeSurvey and imported it to IBM's SPSS 22 (IBM, 2013) for statistical analyses. Our intention was not to draw comparisons here, but to calculate the required sample sizes for our main experiment. Therefore, we do not present inferential statistics here, but rather the descriptive statistics and SPSS's partial eta-squared (η_p^2) statistic as a measure of effect size for our dependant variable.

Session length

For session length we first analyse the data from all three conditions, and again excluding the uniform results from the shutdown conditions Table 8.3. (Incidentally, this data behaves in a similar fashion to the main experiment, with the fatigue condition producing the longest sessions by nearly thirty minutes.) However, the ANOVA performed between these two conditions did not indicate that this effect is significant $F(1, 7) = 7.023$, $p < 0.033$, $\alpha = 0.05$ (two-tailed test, so critical $\alpha = 0.025$).

Table 8.3 Descriptive statistics session length. Time measured in minutes. Shows the effect size calculation both with and without the shutdown condition.

Condition	Number	Mean	SD	η_p^2
Fatigue	10	134.3	27.32	0.794
Control		105.1	28.86	
Shutdown		60	0	
Excluding shutdown				
Fatigue	10	134.3	27.32	0.501
Control		105.1	28.86	

Flow

The Flow Short Scale (FSS) data for one of the participants in the CFS is missing from the raw. Therefore, there are only nine sets of data (three from each group) contributing to the values seen in Table 8.4 p. 134. For all factors in the FSS, we test our hypothesis using a repeated measures ANOVA test. For flow, the shutdown condition produces the highest mean, followed by control and fatigue respectively. Fatigue produces the highest importance scores.

Affect

We performed the repeated measures ANOVA test on the data from the affect schedule. Participants recorded the highest positive-affect (PA) after the control sessions followed closely by the shutdown sessions. They displayed the least PA after the fatigue condition. The test did not identify any significant effects, $F(2, 14) = 2.911$, $p < 0.088$, $\alpha = 0.05$ (two-tailed test, so critical $\alpha = 0.025$). The effect for this dimension is likely to be large.

Table 8.4 Descriptive statistics for the Flow Short Scale. Scores for Flow and Importance are on a scale of 1-7. The midpoint of the scale is four. Scores for Demand, Skills and Perceived fit are on a scale of 1-9. The midpoint of the scale is 5

Factor	Condition	Number	Mean	SD	η_p^2
Flow	Fatigue	9	4.81	0.79	0.165
	Control		4.93	0.97	
	Shutdown		5.23	0.94	
Importance	Fatigue	9	4.29	1.82	0.339
	Control		3.33	1.34	
	Shutdown		3.55	1.82	
Demand	Fatigue	9	2.66	1.58	0.181
	Control		2.77	1.09	
	Shutdown		3.44	1.94	
Skills	Fatigue	9	6.33	1.41	0.157
	Control		7.11	0.60	
	Shutdown		6.55	1.81	
	Fatigue	9	4.88	1.05	0.147
	Control		4.44	1.01	
	Shutdown		4.22	1.64	

Table 8.5 Descriptive statistics for the International Positive And Negative Affect Schedule Short Form. Scores are on a scale of 5-25. The midpoint of the scale is thirteen.

Factor	Condition	Number	Mean	SD	η_p^2
Positive-affect	Fatigue	10	14.00	2.62	0.294
	Control		17.00	2.54	
	Shutdown		16.40	4.79	
Negative-affect	Fatigue	10	6.40	1.71	0.420
	Control		5.80	0.63	
	Shutdown		8.00	2.53	

Participants displayed the least negative-affect (NA) after plying the control condition; the mean is only points stronger than no NA. The fatigue condition is associated with the next highest NA, and the shutdown condition shows the highest (Table 8.5). (Incidentally, the test did find a significant effect $F(2, 14) = 5.077$, $p < 0.022$, $\alpha = 0.05$). Pairwise comparisons indicate that the effect is between the control and shutdown conditions.

H.5 Discussion

Sample size calculations

Using this data, we estimate the sample-size required for adequate power in all ANOVA test we planned to run. We used G*Power (Faul, Erdfelder, Lang, & Buchner, 2007), a power analysis tool to make the estimations. Specifically, we used the *F tests* test family and selected *ANOVA: Repeated measures, between factors* for the statistical test because we want to see if effects exist between the session length factors. We then selected *A priori: Compute same size – given α, power, and effect size*. We opted to specify the effect size *as in SPSS,* because we use *SPSS* for the research; we *determined* our *effect sizes f* directly from the *partial-eta² (η_p^2)* values calculated by SPSS.

See Table 8.6 for the specific values used in each sample size calculation, and Figure 8.8 (p. 136) for a screenshot of G*Power with the relevant selections. We chose α = 0.05 and Power of 0.95 (0.8 for flow) as acceptable error rates because the risks associated with Type I and Type II errors are low for our research. Our pilot data describes the desirable experiences of moderate to high positive-affect and flow state, and low negative-affect for most participants in all conditions. The increased session length observed in the fatigue condition is relatively small compared to the marathon session lengths commonly associated with problematic gaming behaviours. We measure three factors (or *groups* according to G*Power) namely the shutdown, fatigue and control conditions. Each participant contributes three *measurements.* The total sample size describes the total number of samples required across all conditions. For example, the ANOVA for positive-affect requires fifteen samples per condition to identify a significant effect of size f(U) = 0.645 with power of 95%. This translates into fifteen individual participants. Table 8.6 shows that the flow variable would require a substantially larger group of participants (N = 43) to satisfy these error rates.

Table 8.6 Values used for calculating required sample size in G*power. We used the F tests: ANOVA: Repeated measures between factors. Values for each dependent variable estimated from pilot-study data. Error probability α is 0.05 and power is 0.95 for all DV expect where we also calculate sample size for flow with power of 0.8. Number of groups and Number of measurements both are 3. Options: Effect size specification as per SPSS. See Figure 8.8 p. 136.

			Effect Size		Total	Participants
Dependent Variable	α	Power	SPSS η_p^2	f(U)[36]	Sample Size	(+15%)
PA	.05	0.95	0.294	0.645	45	15 (18)
NA	.05	0.95	0.420	0.851	30	10 (12)
FSS	0.5	0.95	0.113	0.357	129	43 (53)
	.05	0.8	0.113	0.357	84	28 (33)
Session length	.05	0.95	0.794	1.96	12	4 (5)
Control & Fatigue	.05	0.95	0.356	1.00	36	12 (14)

Our repeated measures design saw participants attending one session per week at the same time each week for three consecutive weeks. This requires substantial time commitment from both participants and the researchers/observers. Repeated measures designs are prone to attrition whereby participants fail to attend the measurement events or drop out of the experiment at rate of 5% per week (Bhaumik et al., 2008) for a total of 15% across the entire period. This suggests that our study requires 53 participants in total. As our resources are relatively limited (especially time) we calculated the required sample size for flow using the lower bound of generally acceptable power = 0.8 (Cohen, 1988). Flow is only a positive experience, and the risk of false negative error in this test is negligible. This suggested we required 28 participants at these error rates, and we would therefore need to recruit at least 33 participants to account the attrition rates

[36] Calculated by G*Power

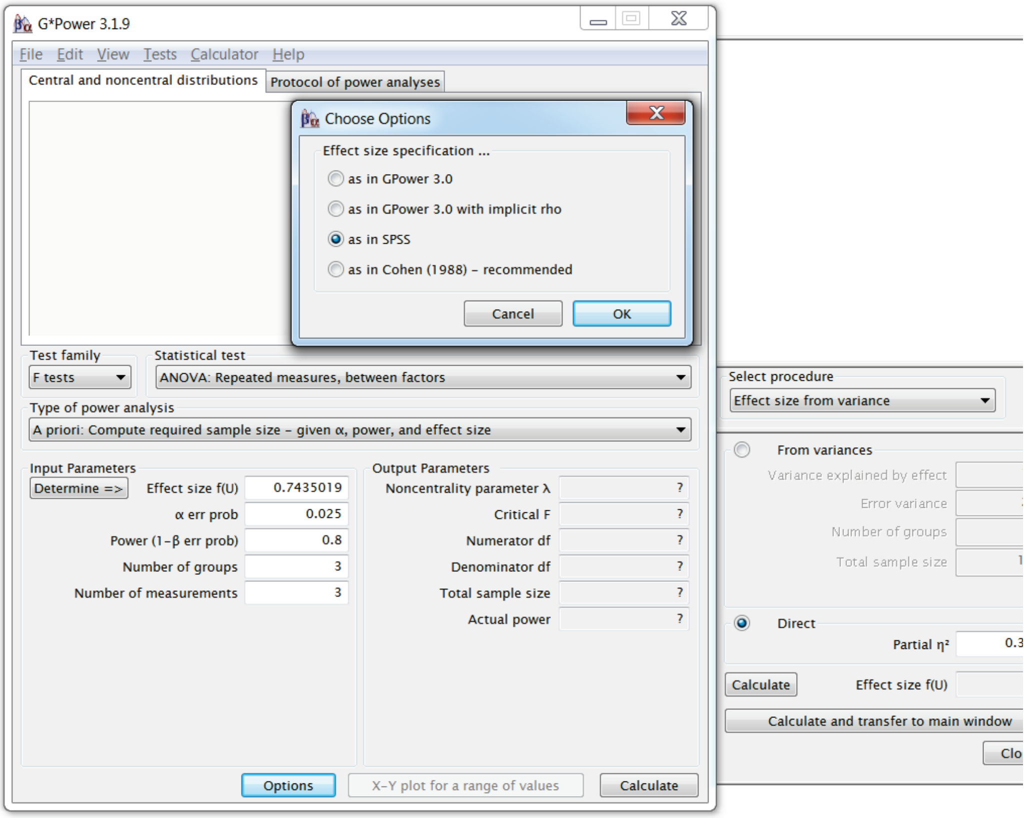

Figure 8.8 Screenshot of sample size calculations in G*power All options and selections made for sample size calculations using G*Power, including the Options setting for Effect size calculations as in SPSS

We did not exhaustively analyse the qualitative information we collected because these would not influence our sample size calculations. One statement inspired an additional measurement outcome. On participant, after their second session which was a shutdown condition, said:

P: "Now all I want to do is play Torchlight II. I might download it again…"

Something they had not considered in the previous fatigue condition, raising the possibility that the condition might affect how soon players choose to return to playing. The notion that a behaviour policy might increase player's desire to return to the game bears thematic similarity to the stories of players using others' national identifications to circumvent the real-name registration policies and continue playing games (Olson, 2016). Thus, we added an item to our battery of questions asking players how soon the wished to return to the game (see Appendix G).

H.6 Conclusion

We accomplished our primary goal with this pilot study: sample size calculation. Although the number of participants required for flow tests are substantially larger than the other dependant variables, the numbers are acceptable.

Appendix I ADDITIONAL STATISTICAL ANALYSES

I.1 Repeated Measures ANOVA: non-sphericity corrections

When data does not show sphericity, we can correct the degrees of freedom to account for this property. For example, when we perform a repeated measures ANOVA in SPSS for our time factor on all 99 data points, Mauchly's test for sphericity reports the statistics seen in Table 8.7. The key statistics significance value, and the epsilon values are the key statistics here. The significance value suggests that if the population does exhibit sphericity, then there is a 2.7% chance that another random sample will produce the deviation in sphericity we observed in our sample. Our data therefore violates the assumption of sphericity and we must use a corrected value for degrees of freedom.

Table 8.7 The statistics produced by Mauchly's test of sphericity for our session length factor.

					Epsilon		
Within-subjects effect	Mauchly's W	Approx. χ^2	df	Sig. (p)	Greenhouse – Geisser	Huynh-Feldt	Lower-bound
Condition	0.780	7.214	2	0.027	0.820	0.918	0.500

Table 8.8 The corrected degrees of freedom (df) values based on Mauchly's W test for sphericity.

Correction	df	F	Sig. (p)	η_p^2	Observed Power
Sphericity assumed	2	63.120	< 0.0005	0.678	1.000
Greenhouse – Geisser	1.639	63.120	<0.0005	0.678	1.000
Huynh-Feldt	1.837	63.120	< 0.0005	0.678	1.000
Lower-bound	1.000	63.120	< 0.0005	0.678	1.000

According to Girden (1992), when Mauchly's test shows significance, then we must look at the Greenhouse – Geisser calculation for epsilon to decide which correction to use. If the value is less than or equal to 0.75, then we use the statistics associated with the Greenhouse – Geisser correction in Table 8.8. If the value of the value is greater than 0.75, we use the statistics associated with the Huynh-Feldt correction. For session length, the Greenhouse – Geisser epsilon is 0.82, therefore we used the Huynh-Feldt correction. In this specific case, the correction makes no difference to the crucial values derived from the test.

Table 8.9 Descriptive stats for session length of each participant group. Shutdown condition excluded because the session is always sixty minutes.

						95% CI for Mean		
		N	Min	Max	Mean	Lower	Upper	SD
Control	Shut-Con-Fat	11	70	163	105.73	87.54	123.92	27.078
	Fat-Shut-Con	11	60	130	94	79.07	108.93	22.222
	Con-Fat-Shut	11	80	170	112	93.51	130.49	27.525
Fatigue	Shut-Con-Fat	11	70	142	117.55	100.92	134.17	24.744
	Fat-Shut-Con	11	82	242	145	111.97	178.03	49.173
	Con-Fat-Shut	11	70	179	121.09	96.16	146.03	37.117

I.2 Session length

We present descriptive statistics of session length for each sequence group in the control and fatigue conditions in Table 8.9 p. 137.

I.3 Flow

Interaction effects

We performed the repeated measures ANOVA for each sub factor of the Flow Short Scale and found no significant interaction effects between participant groups. Effects for our sample are small, moderate or large. Observed power is especially low, but unsurprising where p-values are greater than 0.5. See Table 8.10.

Table 8.10 Inferential statistics for the Flow Short Scale: participant-group interaction effects.

Component	F	df	Sig. (p)	η_p^2	Observed power
Flow	1.730	2	0.194	0.103	0.334
Adsorption	2.098	2	0.140	0.123	0.397
Performance	1.127	2	0.337	0.070	0.229
Importance	0.531	2	0.593	0.034	0.129
Demand	0.087	2	0.917	0.006	0.062
Skill	0.010	2	0.990	0.001	0.051
Fit	0.101	2	0.905	0.007	0.064

I.4 Affect

Negative-affect – individual emotions

The results of the Friedman tests performed for the individual emotion in the I-PANAS-SF (see Table 8.11). We found a difference only for *upset*.

Table 8.11 Results of Friedman's test for the individual emotions that make up negative-affect.

Dimension	N	χ^2	df	Sig. (p)
Upset	99	28.659	2	< 0.0005
Hostile		5.815		0.550
Ashamed		4.294		0.117
Nervous		0.375		0.829
Afraid		2.182		0.336

Comparisons with Thompson

Our positive-affect and negative-affect data appears normally distributed when all conditions are considered. We performed a one-sample t-test against the values reported in Thompson (2007). We found statistically significant differences for all conditions in both PA and NA. We present the statistics in Table 8.12. See Section 5.5 for discussion.

Table 8.12 One-sample t-tests comparing our affect results with Thompson's (2007) means. Thompson's reported positive-affect (PA) mean is 19.15, and negative-affect (NA) mean is 12.73 (2007).

	Condition	Mean (M)	t	df	Sig. (p)	Mean Difference (M – 19.15)	95% CI of the difference Lower	Upper
	Fatigue	14.85	-5.793	32	< 0.0005	-4.302	-5.81	-2.79
PA	Control	15.61	-4.902	32	< 0.0005	-3.544	-5.02	-2.07
	Shutdown	15.30	-5.186	32	< 0.0005	-3.847	-5.36	-2.34
	Condition	Mean (M)	t	df	Sig. (p)	Mean Difference (M – 12.73)	95% CI of the difference Lower	Upper
	Fatigue	7.03	-13.403	32	< 0.0005	-5.700	-6.57	-4.83
NA	Control	5.87	-36.971	32	< 0.0005	-6.912	-7.29	-6.53
	Shutdown	7.42	-13.714	32	< 0.0005	-5.306	-6.09	-4.52

Appendix J RAW REASONS FOR STOPPING

We summarized each response by identifying the key words or key phrases within. This process was trivial for many responses, as they contained the same words, including *tired* and *hungry,* and phrases conveying the idea that some other activity required the attention of participants. Furthermore, we identified four common themes from all the responses. Where responses fell into two categories, we prioritized the reason that effected the participant most immediately, often discomfort. Where we had to choose we have provided justification

Table 8.13 Themes, ideas, and description of categories for all participant's reasons for stopping. Derived from key words, phrases and ideas in the raw responses (see Table 8.14, Table 8.15, and Table 8.16).

Category	Theme	Description	Key words, phrases and ideas
Other Activity	I want to carry on playing but I have another commitment.	Another activity of social consequence requires the participant's attention. They ended their session to attend to it.	Meeting, lectures, study, work, work to do, I need to
Satisfied	I completed all pressing goals	Participants accomplished their game related goals, or overcame an obstacle and felt that continuing would be as starting another activity.	Finished, Level up finished quest/boss, good place to stop.
Discomfort	I have enjoyed playing, but I would rather not continue right now	Participants realize that continued play would produce - or continue to produce - undesirable experiences of a personally psychological or physiological nature.	Tired, hungry, sore, bored, difficult, Frustrating
Discomfort: Fatigue system noticed	I do not wish to continue while the game is fatigued.	Sub-category of discomfort, where participants explicitly identify the fatigue system explicitly as the source of discomfort.	Gold, XP, Experience, value, progress too slow

Table 8.14 Each participant's reason for hiatus in the control condition. We present the category we placed it in, and why we placed it there. Responses verbatim.

Response	Category	Utterance
Finished some big quest where I died a lot	Satisfied	Goal accomplished
Got my level up	Satisfied	Goal accomplished
I felt I had done enough and I had finished my current quests.	Satisfied	Goal accomplished
I finished what i had set out to do.	Satisfied	Goal accomplished
had other things that needed to be done, such as studies	Other activity	"Other things" to do
Had to go back to games project..	Other activity	"Had to"
Had to meet someone, otherwise I would have continued since I was enjoying myself	Other activity	"Had to"
I have a lot of other work to do, so I had to cut my play time short so that I can finish it first. Once I'm done with my work I might play again.	Other activity	"Had to"
I have so much work to do.	Other activity	"I have … to"
I have to study, unfortunately.	Other activity	"I have to"
I need to get to lectures soon and I didn't want to cut it too fine, so I decided I would stop at 11:15 and that's when I stopped.	Other activity	"I need to"
I stopped playing because I have other work to do.	Other activity	"I have … to"
Lost track of time and realised that I had other things that needed to be done so I stopped playing	Other activity	"other things … needed to be done"
My lift arrived	Other activity	Another person waiting on them
Remembering other responsibilities intuitively made me stop playing (kind of like an internal alarm clock). Although I would have loved to keep playing all day.	Other activity	"Other responsibilities"
Unfortunately have to go home and do some work.	Other activity	"Have to"
It's getting late. I have things to do	Other activity	"Have to"
I have to leave now to beat traffic	Other activity	"Have to"
Got Tired, I had completed the mission I intended to and was wrecking ALL the scrubs [game related term for weak opponents] easy.	Discomfort	Tiredness precipitates exit point location
hungry	Discomfort	Hunger

Table 8.14 Each participant's reason for hiatus in the control condition. We present the category we placed it in, and why we placed it there. Responses verbatim.

Response	Category	Utterance
I became hungry. If I was playing at home, I would have paused, found lunch, and carried on. The exact time of my halting was because my character died, and I usually quit only after I am back in town (so I can start afresh each game)	Discomfort	Hunger
I felt tired.	Discomfort	Tiredness
I got hungry	Discomfort	Hunger
I had played for 2+ hours and was tired. I had completed all the quests prior to the last dungeon.	Discomfort	Tiredness precipitates exit point location
I was feeling a bit tired and just finished a big quest so I decided to give the game a rest	Discomfort	Tiredness noticed after goal reached
I was hungry, and I wanted a break as my eyes were hurting. I only stopped after I had finished my current major quest and cleaned up a pending side quest.	Discomfort	Hunger and pain precipitates exit point location
I was tired and had reached a good save point.	Discomfort	Tiredness coincides with exit point location
Slightly worn after 1h30 of concentration. Also, have other responsibilities to attend to. Could easily continue after a 20 minute break.	Discomfort	Tiredness precipitates exit point location
Started getting hungry. Liked having the control to decide when to stop. Last time the game just quit and this made me anxious as I was not sure if my progress was saved. I also felt a bit disoriented afterwards, the last time. This time not so much.	Discomfort	Hunger
tired because it was repetitive actions . stating to get a bit predictable. and also that i stated to lose health in the game too much , and ran out of money/coins more often.	Discomfort	Tiredness Boredom difficulty
Tired.. Also thinking about my girlfriend - I have plans to see her after this session. RSI [repetitive strain injury] problem keeps me from	Discomfort	Tiredness Pain Longing?

Table 8.14 Each participant's reason for hiatus in the control condition. We present the category we placed it in, and why we placed it there. Responses verbatim.

Response	Category	Utterance
wanting to use a computer for too long.		
Got a bit bored.	Discomfort	Boredom
GOT DEFEATED BY MY OPPONENT (participant complained of difficulty towards the end)	Discomfort	Difficulty

Table 8.15 Each participant's reason for hiatus in the fatigue condition. We present the category we placed it in, and why we placed it there. Responses verbatim.

Response	Category	Utterance
After completing a challenging task (defeating the big ogre thing at the end of the first dungeon), the idea of dealing with other tasks in the game; following and executing a plan; keeping track of my progress, all seemed like too much effort.	Satisfied	Predicting tiredness and boredom, Pre-emptive exit at satisfactory point
Finished the "act"	Satisfied	Goal accomplished
I had just finished a major quest and collected the reward, so I felt like it was a good place to stop. I was also feeling a bit tired, as I had been playing for about two hours. I do also need to get to lectures fairly soon, so I didn't want to start on the next major section - especially seeing as though this was my last session.	Satisfied	Goal accomplished, was a bit tired, other activity later, not compelled to commit to more. The tone is "I could have done more, but I didn't need to."
Played for quite a while. Felt like it was a good place to stop.	Satisfied	No discomfort or other activity mentioned. Not compelled to continue. Exit point "felt good."
Reached goal - decided to stop after boss was killed	Satisfied	Goal accomplished
had an appointment	Other activity	Appointment
I have other things to do	Other activity	"I have … to"
I need to do other things (study).	Other activity	"I need to"
It was getting late and i have class tomorrow,	Other activity	"I have"
My lift arrived	Other activity	Another person waiting on them
Time, I had other obligations (also felt like enough)	Other activity	"I had … obligations"
I stopped playing because I want to go have a beer. If my schedule was free, I would have probably kept on playing. I made sure that I	Other activity	Other activity takes priority

Table 8.15 Each participant's reason for hiatus in the fatigue condition. We present the category we placed it in, and why we placed it there. Responses verbatim.

Response	Category	Utterance
finished the quest I was on at the time before stopping.		
Got tired of playing. Probably played a bit longer than I would have otherwise because trying to miss traffic.	Discomfort	Tiredness, other activity not time crucial
I got tired. I started making more mistakes because I was fatigued, and that became frustrating	Discomfort	Tiredness
i have to go and eat lunch ,i didn't ate my breakfast	Discomfort	Hunger
Game's gettting grindy	Discomfort	Boredom
I did not know how long I had been playing and was somewhat aware that there might be another person who might come to play after me. I did not want to take over their session.	Discomfort	Apprehension, uncomfortable with continuing
I didn't feel like continuing	Discomfort	Boredom implied
It got to a point I respawn and then I just die.	Discomfort	Difficulty
Tired. Getting a Bit bored . it was becoming repetitive and too predictable...	Discomfort	Tiredness Boredom
Tired	Discomfort	Tiredness
Getting quite repetitive. Every item I get from a quest I can just buy a better one	Discomfort	Boredom
Although I realised the gold's value decreased, I don't care that much about that (typically with games like this you find/earn the best items, so I had no real use for the gold). However, I soon realised that my experience was no longer increasing in a noticeable manner. I thought it was pointless to continue, as I play such games mainly to advance & learn new character skills. I actually felt like the game was broken, and that a restart might solve it.	Discomfort: Fatigue noticed	Gold and experience "no longer increasing"
Deminishing returns and not enough health potion droppings to sustain combat... too much effort to go back to town to buy more supplies	Discomfort: Fatigue noticed	"Deminishing returns"
Experience gain and item drop chance was too low for it to feel rewarding	Discomfort: Fatigue noticed	"[loot] drop too low"

Table 8.15 Each participant's reason for hiatus in the fatigue condition. We present the category we placed it in, and why we placed it there. Responses verbatim.

Response	Category	Utterance
I couldn't keep up once I stopped leveling up, I became quite difficult to battle with the bosses and champions. I ran away most of the time.	Discomfort: Fatigue noticed	Difficulty increased "once I stopped leveling up"
I knew I had only so long before the mod would kick in and I would begin losing out on exp and gold. Having played this game before, I knew what I could accomplish within that time and conquered specific areas with that in mind.	Discomfort: Fatigue noticed	"I knew I would begin losing out on exp and gold"
Leveling up became so slow that I could not compete against enemies. Took forever to kill people. No feeling of progress. Also I got a bit bored and slightly hungry as was a long continuous period.	Discomfort: Fatigue noticed	"Leveling up became so slow"
Too difficult. XP rate left me on level 17 when the dungeons were for 25+	Discomfort: Fatigue noticed	Difficulty due to fatigue – level differences
XP gains became absolutely nothing and there was no point in playing anymore	Discomfort: Fatigue noticed	"XP gains became absolutely nothing"
The game is fun but I haven't leveld up in an hour	Discomfort: Fatigue noticed	"I haven't leveld up in an hour"
violent	Unknown	Excluded
	Unknown	Excluded

Table 8.16 Each participant's reason for hiatus in the shutdown condition.

Response
the game exited on its own accord.
The game automatically closed.
I was forced to.
The game cut out for the purposes of the test.
The game cut out after an hour as expected
game closed
Because the game stopped while I was playing
it stopped itself
The time ran out.
Was kicked out :'(
Time ran out and I didn't even notice
The game stopped
the 1hr time limit expired
Game dropped after 1 hour :/
Well, it was not by choice. I played my hour through then the game closed
The game kicked me out.
IT SHUTDOWN.
I was made to stop
The game abruptly ended
It exited automatically
time out
The game just abruptly ended. :(
Game time expired ;(
Because I was kicked off the game. At least I started Act II and leveled up one more time (level 18 or 19) just before I was kicked.
Experimentor cut me off after an hour.. dick.
The gae exited automatically
Because I was forced to stopped.
Because time ran out
Hour period ended.
1h timer ran out
It shut me out at 1 hour
Game over
The game stopped in the middle of a battle!