

# The Relationship Between Subjective Questionnaire Scores and Metacritic Scores

Christie Pei-hang Lau

**Project report submitted in part fulfilment of the requirements for the degree of Master of Science (Human-Computer Interaction with Ergonomics) in the Faculty of Brain Sciences, University College London, 2015.**

**NOTE BY THE UNIVERSITY**

**This project report is submitted as an examination paper. No responsibility can be held by London University for the accuracy or completeness of the material therein.**

# ACKNOWLEDGEMENT

First of all, I would like to thank my supervisor, Dr Charlene Jennett, for her guidance, help, patience and encouragement throughout the whole project. I would also like to thank Dr Jo Iacovides and Dr Anna Cox for their insightful advices.

I am grateful to Cyril Rebetez in Sony Computer Entertainment Europe Limited for his help on participant recruitment, helpful advice on games selection and statistical analysis, and for his kindness of lending the Sony Playtest Lab and resources for my study.

Thanks to all the participants who dedicated their time in my experiment.

Finally, I would like to thank my family and friends for all their help and support during this project and throughout this course.

# ABSTRACT

This study aimed to examine how measures of player experience used in videogame research relate to Metacritic Professional scores, and how well these measures are able to segregate games of different Metacritic scores. 39 participants played five First Person Shooter (FPS) games and their responses to the Immersive Experience Questionnaire (IEQ), the Player Experience of Need Satisfaction (PENS), the Game Experience Questionnaire (GEQ) and the industry questionnaire were recorded.

Across all four questionnaires, differences in game component scores were found with Control in IEQ, Intuitive Controls and Autonomy in PENS, and Challenge and Competence in GEQ, which are related to player's ability to control and navigate within the game. Correlations among the data indicate a mere overlap between some of the player experience constructs and the factors informing Metacritic scores. However, only some of the questionnaires were able to segregate games of different Metacritic scores. The data provide clear evidence that Metacritic scores do not reflect the full complexity of player experience, nor the player experience components defined by academic researchers matches completely how the professional game critics contributing to Metacritic scores review a game. Further research is needed to bridge the existing gap between research and commercial practices in the game industry, and to develop a set of consistent and standardised tool in evaluating game quality.

# CONTENTS

ACKNOWLEDGEMENT

ABSTRACT

CONTENTS

Chapter 1. Introduction

Chapter 2. Literature Review

Player Experience Evaluation

Product quality evaluation

The dichotomy of GUR and business perspectives

Focus on Questionnaires

Popular Videogames Questionnaires

Research questions

Chapter 3. Method

Participants

Experimental Design

Materials and Apparatus

Procedures

Pilot study

Chapter 4. Results

Chapter 5. Discussion

Strengths in current research

Limitations and Confoundings

Current restrictions in studying player experience

Limitations of questionnaires as a research methodology

Future research and implication

Chapter 6. Conclusion

Reference

Appendix 1: Initial questionnaire

Appendix 2: Pre-game questionnaire

Appendix 3: Industry Questionnaire

Appendix 4: Immersive Experience Questionnaire (IEQ)

Appendix 5: Game Experience Questionnaire (GEQ)

Appendix 6: Player Experience of Need Satisfaction Questionnaire(PENS)

Appendix 7: Post-game questionnaire

Appendix 8: Photos of Lab set-up

Appendix 9: Instruction sheet for each game

Appendix 10: Information sheet for participants

Appendix 11: Informed consent form

Appendix 12: Latin Square of Gameplay sequence

# Chapter 1. Introduction

The game industry is growing fast. The games business is expected to grow about 59 percent to \$100 billion worldwide by 2019, according to market researcher DFC Intelligence (2015). Since higher-quality games tend to sell better, game designers and developers are increasingly looking for ways to improve their games, in order to grow their market shares. Assessment of game quality is a window for commercial game publishers to understand how to create better games, and this has led to the rise of a new game industry field called game user research (GUR). GUR researchers adopt scientific user-testing approaches backed by academic domains of psychology and human-computer interaction (Nacke, 2015), to maintain the competitive edge of aspects of their games in this expanding market.

Within the GUR and the academic domains of human factors, human-computer interaction, and social psychology, a blooming number of robust scientific researches have discovered the positive and negative impacts of video games (Anderson et al., 2010; Jones et al., 2014). This contributed to the growing number of game quality assessment methodologies in order to study these impacts. These psychometric scales or questionnaires attempt to unveil the complexity of player experience (PX) with a multidimensional approach.

While the GUR community adopts these psychometric scales or questionnaires developed in academia that emphasise the multidimensional properties of PX, the majority of industry often assess video games based on product quality evaluation. Its results, often in the form of a single numerical score, are important indicators internally for design iterations and sales forecasting, and externally for marketing and customer references. The duality of GUR and business perspectives has inspired this thesis.

The exploration into the dichotomy of GUR and business perspectives hints one of the main challenges the game industry is facing. There is a lack of a consistent and precise set of methodologies and tools which allows the measurement of game quality in a sensitive, reliable and valid manner. Finding the right data source for different measures can be tricky. Validation of existing tools has then inspired another research interest in the topic.

Metacritics has become an important indicator of a game's success (Swain, 2008), and questionnaire is a widely adopted subjective experience evaluation tool in the game research industry as it is cheap and quick to implement (Adams & Cox, 2008). Combining the interests in the dichotomy of GUR and business perspectives and validation of existing tools, this thesis aims to investigate into the relation between Metacritic Scores and questionnaire as a tool.

Past research has investigated the relation between Metacritic Professional Scores and two of the more frequently used subjective player experience questionnaires. Correlations have been found between Metacritic Scores and components in the questionnaires, indicating an overlap between the factors influencing Metacritic scores and the player experience dimensions. (Johnson et al., 2014) However, this work only included evaluations of players' current favourite game, and could not account for the full range of variability in an individual's player experience due to the between subject design. In this thesis we are interested in comparing more subjective player experience questionnaires, investigating whether these questionnaire scores can segregate games of Metacritic Scores in the same group of players.

We conducted an experiment to explore the relationship between Metacritic Professional Scores with component scores of four different questionnaire, three arises from academic research (Game Experience Questionnaire, Player Experience of Need Satisfaction questionnaire, and Immersive Experience Questionnaire) and one from the industry (Industry questionnaire). The findings revealed that the Metacritic scores do not reflect the full complexity of player experience, and there is a mismatch between how the academic researchers and professional game critics consider player experience.

This thesis is divided into chapters:

- Chapter 2 discusses the literature related to this study and explains the reason of choosing Metacritic Score and the questionnaires.
- Chapter 3 describes the method of the experiment to investigate the research questions.
- Chapter 4 displayed results and findings from the conducted experiment.
- Chapter 5 provides a general discussion on the findings and its implications and limitations.
- The last chapter, chapter 6, summarises the key ideas in this study.

## Chapter 2. Literature Review

This chapter will start with a brief review of existing methodologies in evaluating game quality, which is branched into player experience evaluation and product quality evaluation. These two distinct game quality approaches, each has its caveats, will be discussed. This gives rise to an exciting research opportunity to bridge the existing gap between game user research (GUR) and commercial practises in the game industry.

During the discussion of using a suitable research methodology, it will become apparent that the lack of a consistent way to measure game quality can be a hindrance to evaluation the subject matter in a scientific manner. Therefore, another research opportunity lies in validating existing tools.

Among the abundance of tools and methods for measuring game quality there are two approaches: player experience evaluation and product quality evaluation. In this review we aim to give a broad perspective of existing methods and tools from both approaches, explaining the chosen methods, followed by defining our research questions.

### Player Experience Evaluation

Player experience (PX) describes the user experience (Hassenzahl and Tractinsky 2006) in the context of play. A lot of concepts have been proposed to investigate PX, such as enjoyment, immersion and flow, (Brown & Cairns, 2004; IJsselsteijn et al., 2007; Jennett et al., 2008; Nacke & Lindley, 2008) but up to now there is no common consensus on how to describe the phenomena of PX. Despite the current debate, there are two fundamental pathways in studying PX, which are the collection of behavioural data (observing players) and attitudinal data (asking players) (Pagulayan et al. 2003). Currently the overwhelming majority of GUR work is qualitative, mostly usability testing (Bernhaupt, 2010). In this section we describe the kinds of player experience evaluations: Telemetry, biometrics, and questionnaires.

#### Telemetry

Telemetry is data logged from servers about how players play games, or about how the game itself responds to player behaviours. The access to objective, real-time data about how people play games, including all actions and interactions taking place, even at a granular level, enables specific game

elements to be evaluated in detail. For example, game researchers would be able to pinpoint that at the time when a tester experienced A, they were doing B in the game, and the PX at that time point can be compared to its respective design goal. Collection of a vast amount of behavioural data from the entire populations of players is possible, but this poses the difficulty in filtering and analysing such a big data set, especially without enough contextual data. There are not enough existing research or knowledge in telemetry-based evaluation, how to merge it with other user research methods, nor to infer PX with such data.

## **Biometrics**

Biometrics commonly refers to technologies that measure human physiological response signals or characteristics, and these measurements are analysed to understand the related psychological effects, e.g., brainwaves, muscle activity, or skin conductance. The use of biometrics is on the rise in the recent years as the cost of sensor technologies continues to drop (Nacke 2013). We might gain useful, but limited, insights for evaluating arousals, excitement, emotions or cognitive workload from the most common physiological response signal measures. Caution is required when interpreting the physiological responses as it is often the case that one physiological response is associated with many psychological effects, i.e. a many-to-one relationship (Cacioppo et al. 2007). Hence a direct causal relationship between a psychological effect and an emotional state is impossible. Another challenge lies in obtaining a clean signal and usable data from different low-cost devices. Signal cleaning procedures and analysis with different physiological sensors have to be handled with care. Nacke (2015) concluded that psychophysiological measures in games should always be used with other measures and data sources to triangulate relationships between biometrics data and dimensions of player experience.

## **Questionnaires**

Questionnaires have also been used frequently to gather information from players, as it taps into players' perceptions of games (Davis, 2005), assesses a gamer's type or demographics (Eysenck et al. 1985; Nacke et al. 2014), and opens up the possibility of collecting a vast amount of self-report data from a large number of players at the same time. These questionnaires often feature Likert-scale (Likert, 1932) rating type of questions. Ideally, a questionnaire aiming at players' value judgements about gameplay moments is delivered to players immediately after a gameplay session, so that the experience is still vivid in memory.

The subjective metrics that arise from questionnaires can be helpful when triangulated with data sources from other approach (Adams & Cox, 2008), such as biometrics or telemetry, examples of objective, behavioural data sources. Once designed and validated, questionnaires are quick and cheap to implement and to obtain quantitative data, which can provide insights into players' feelings and attitudes fed by statistical analysis.

Game publishers, such as Sony Computer Entertainment, have developed internal questionnaires. However these questionnaires cannot reliably segregate games of different quality and sometimes receive very similar ratings due to central tendency bias. In addition, the small sample size and the fact that the games are not finished when they are tested poses challenges on the creation of an effective and reliable questionnaire. More popular questionnaire sets in game research community are focused on different dimensions of player experience, which will be discussed further in later sections.

By explicitly tapping into players' attitudes and opinions, researchers, both in academia and industry, hope to discover more about the subjective experience of gameplay and ensure that it corresponds to the game designers' intent (Nacke, 2015).

## **Product quality evaluation**

Evaluation of game quality and the numerical ratings that arose from it are important indicators internally for design iterations and sales forecasting, and externally for marketing and customer references. Marketing, sales and financial functions in the game industry often assess video games based on product quality evaluation. Currently there are a few ways of measuring game quality as a product, with its results often a single numerical score, for example, market research, game reviews, Metacritic professional scores and internal game testing, to cite a few. Below each of these methods and its limitations are discussed.

### **Market Research**

Market research provides insights to game publishers in predicting responses of a to-be-released product in a certain market, and such prediction has high importance in resources allocation within an organisation. (Malhotra, 2014; Reichheld, 2003) concluded that measuring customer satisfaction was not a strong enough indicator of predicting purchase and referral behaviours, which are often tied

to important revenue driving streams in many companies. The Net Promoter Score was then developed, and has become a widely adopted metric in market research in measuring customer loyalty, which could consistently predict short-term purchase and referral behaviours (Reichheld, 2006). The score is produced by a 11-point Likert Scale item, “On a scale from 0-10, how likely are you to recommend this game to a friend or colleague?”. There is a significant correlation between customers’ tendency to recommend a company to their rate of purchase and referral behaviors (Satmetrix. 2004). However, Keiningham et al (2007) raised conflicting results, critique on the question does not provide an as strong business growth predictor as it claimed in the original study.

Although the Net Promoter Score can be useful in predicting customer loyalty in game purchase, it aims more at prospecting audience appeal rather than assessing the game's quality as in a development cycle or as a finished product, and how it is related to actual gameplay experience is not known.

## **Game reviews**

Game reviewers outline and appraise their own subjective player experience, and many of them produce reviews on the internet, in forms of blog articles and/or scores. Livingston (2011) found that player experience can be influenced by reviews when read before playing a game for the first time. The rise of both professional and amateur videogame reviewers, together with the impact brought by fellow players who read the reviews, has bred services which collect reviews from different sources.

Metacritic.com is a popular aggregator website that takes reviews of entertainment products from many sources and normalizes the rating from each into a 100-point scale. A Metacritic User Score is the unweighted average of all reviewer scores. Nacke (2014) suggested that the Metacritic user scores might be composed of disproportionately extremely negative and extremely positive opinions about a particular game. This might be because those who feel less strong to the game might be less likely to devote the time to rate it compare to those who feel strongly about it.

## **Metacritic Professional Score**

Alongside its user score, the Metacritic website produces a Metacritic Professional Score, which is calculated by collecting a number of professional reviews (from sources that review videogames), converting each of the review scores into a score out of 100 (e.g. B+ becomes a score of 83, 3 stars in a

4-star scale equates to a 75). Using a system of weighted averaging, in which more prominent critics, reviewers and publications have stronger influence according to their quality and overall stature, Metacritic.com produces a final aggregated Metacritic professional review score- “Metascore” (Metacritic.com, 2015). Metacritic.com did not disclose the weighting among its sources to the public. This system is valuable because the opinions of many reviewers will average out and thus not be susceptible to the personal tastes of a single reviewer.

Professional game critics review game titles based on a lot more factors other than the actual play experience. For example, reception of the previous title under the same franchise, competition on the market, interest for the certain genre or topics on the market at the time, etc (Bernbeck, 2015).

By analyzing games based on Metacritic scores, a relationship between a high Metacritic score and high sales is revealed. A study by EmSense, Inc. shows that each 5-point increase on Metacritic score correlates to an approximately 50 percent increase in sales, which means that games that scored 75 on Metacritic.com bring about average sales of \$5 million and games that scored 90 accomplish average sales of nearly \$50 million. (Note that these data exclude games based on movie licenses). (EmSense, Inc., 2008). Another study found that a one unit increase in the Metacritic score is found to increase unit sales by approximately 1.5%. Thus, a 10% increase in the review score would typically increase unit sales by around 15%. (Cox, 2014) Despite the results not agreeing with each other as to how strongly does the Metacritic score relate to an increase in sales, it is obvious that review scores promote sales of games as it provides an important guide to customers on purchasing of game titles. It is worth noting that although a correlation between Metacritic score and sales, marketing spend also has a huge influence on sales (Berry & Linoff, 2014). Therefore caution has to be taken to use sales figure to evaluate game quality. However, little is known about how much aggregated reviewer scores reflect the actual player experience (Everiss, 2008, Cox, 2013).

## **The dichotomy of GUR and business perspectives**

A remarkable distinction between the academics and commercial approaches is that the former usually recognises the multidimensional construct of player experience (PX), backed by psychometric measures of PX. It also recognises individual differences, that players might enjoy different dimensions of the same game. Many of the popular game questionnaires address multidimensionality by producing results in forms of component scores, to acknowledge a videogame can have different levels of performance across its dimensions, i.e. the game might be better in some aspects and worst in others. However, it is relatively difficult to understand the multidimensional characteristics of PX

and the component scores by a wide range of people from other disciplines (human computer interaction, social psychology, human factors) who work on the same project teams.

The commercial approach, on the other hand, often relies on a single score as a measure of quality and success. Using a single score is intuitive in many respects and it allows for relative judgements, and it has become a common practice when a game is reviewed at different stages of a game production. However, reducing a videogame to a single score as a measure of quality and success would be seen as “counterintuitive” by academics in the GUR community (Johnson et al, 2014) as this masks complexities of PX.

Derived by analyzing Metascores, Swain (2008) listed characteristics that high-scoring games tend to have. Large in scope—20 hours of content, variety of player choice/activity, high replayability, top quality visuals and sound and responsive and easy controls are ranked as the top five important elements. However following strictly to the listed items does not always lead to high metacritics score. In fact, Swain pointed out that “Gameplay undifferentiated from similar titles” is a top contributing factor to low metacritics scores. Although Swain did not give an example on the factor to low scores, this list has its value as a guideline or checklist to game developers of what makes a game successful, but the relationship between these listed characteristics and the weight of each item in the normalised metacritics score, is not known.

Since popular game questionnaires emphasis on presenting the complexities of PX and unveiling its different dimensions (e.g. variety of player choice/activity, or autonomy in games), would component-based game questionnaires predict metacritic scores? To limit the scope of this research, only the potential links between component scores from questionnaires and Metacritic Professional score will be studied.

## **Focus on Questionnaires**

Questionnaires are chosen as the focus of this research as they can directly measure the subjective PX. This research method is widely used in the majority of academic studies, cheap and easy to deploy, and offers a standardised test for quantifying a particular aspect of PX (Adams & Cox, 2008). Questionnaires also give an opportunity for players to convey their subjective experience, thoughts and feelings, within the criterion set by the questions. The questions, if well written, can guide players to recall and consider specific facets of their subjective experience. Another benefit that standardised questionnaires bring is that they have high reliability when a large number of people are available for testing (Nacke,2015) due to a larger effect size. Best of all, written questionnaires can provide

consistency by ensuring all participants have thought through the same specific concepts guided by the questions.

Alongside the advantages this research method provides, the use of questionnaires is closer in concept to the nature of game review scores, and the sources of the aggregated Metacritic Professional score, for its subjectivity. In essence, game critics summarise and quantify their subjective experience, which is a similar psychological process as players fill in a questionnaire - the questionnaire is a self-reported experience of the players. It is safe to assume that the review scores are informed directly by the quality of a reviewer's PX, but there is little research on the effects of different dimensions of PX on the review process of a game. In contrast, popular game questionnaires can measure PX in specific dimensions of the subjective experience, which will be further elaborated below. Since the Metacritic Score arises from subjective reviews, it is a key investigation interest in this research, and using questionnaire which quantify subjective PX is an appropriate choice. The component scores featured in questionnaires could help identify potential influences of different dimensions of PX on the professional review process of a game.

## **Popular Videogames Questionnaires**

The experience of playing games have been described by terms such as flow, presence and immersion, to cite a few (Brown & Cairns, 2004). The player experience, the experience of being entertained through games, or the user experience of games, is not well understood yet from a psychological perspective (Wyeth et al, 2012). Currently there is an emerging number of research done on different concepts of game experience, such as immersion (Brown and Cairns, 2004; Jennett et al., 2008). It is difficult to decide upon which psychological concepts are best suited to study the subjective experience of play.

This leads researchers to develop different questionnaires in order to validate the concepts of their research interests, attracted by the above mentioned benefits of questionnaire as a research tool. This phenomenon, however, is challenging for new researchers because of the proliferation of questionnaires available.

Which questionnaires are measuring what aspects of experience is not clear. Researchers might not be familiar with precise details of each theory, the key differences between each of them, or the overlaps between the concepts or central components which the questionnaires are developed from. Different

questionnaires are used in each research, with different games, settings and subjects being tested, making it hard to compare the effectiveness of each questionnaire. (Nordin et al., 2014).

The Gaming Experience Questionnaire (GEQ), Player Experience of Need Satisfaction Questionnaire (PENS) and Immersive Experience Questionnaire (IEQ), are prominent examples of questionnaires set up in a similar way in order to evaluate player experience. The main problem is that all these questionnaires have the same inspiration: flow theory (IJsselsteijn et al., 2007; Jennett et al., 2008; Ryan et al., 2006). The theory aggregates everything that a pleasant task / experience should account for: immersion, challenge, sense of achievement, etc. Currently there is no research to test whether they are all measuring the same experience. Nordin et al. (2014) sets out the need for positioning the various questionnaires in relation to each other. They argue that further investigation on these questionnaires is needed to produce better quality questionnaires and reduce confusion amongst player experience researchers.

Given that each questionnaire has its own interpretation of flow theory's dimensions, another research interest, alongside the relationship between component scores and Metacritic scores, is to assess whether users rate each of these questionnaires differently (for instance, GEQ, IEQ and PENS have different components for immersion). This is to investigate whether self-reported questionnaires can reliably measure differences in terms of quality of player experience.

## **Game Experience Questionnaire (GEQ)**

'Characterising and Measuring User Experiences in Digital Games' by IJsselsteijn et al. (2007) acknowledge that it is not an easy task to adequately describe and measure the gaming experience. Yet it is extremely important to do so for the purpose of testing and evaluating games.

Game Experience Questionnaire (GEQ) was designed to play test games and assess seven aspects of the subjective experience of play, which are: Positive Affect (feeling of positive emotions), Negative Affect (feeling of negative emotions), Frustration or Tension (irritation arises from negative experiences induced by gameplay), Flow (sense of acting in the game), Challenge (experience of being tested), Immersion (perceived absorption), and Competence (perceived efficacy playing the game). The GEQ's psychometric properties are questionable because the preliminary validation work (leading to the creation of the scale) has never been published, yet the questionnaire set has become a popular tool in the field since developed.

## **Player Experience of Need Satisfaction (PENS)**

Przybylski, Rigby and Ryan (2010) applied an established psychological theory – self-determination theory (SDT) – to video game player motivations. Based on SDT and other relevant theories (e.g., presence), Przybylski and colleagues developed the Player Experience of Need Satisfaction (PENS) measure. This questionnaire assesses player experience in the dimensions: Competence (feeling of capable in the game), Autonomy (perceived flexibility and self-sufficiency), Intuitive Controls (perception of the in-game controls) and Presence/Immersion (perceived existence in the game world, as opposed to experiencing oneself as a person outside the game, manipulating controls or characters).

The initial validation of the PENS was performed in experimental contexts with participants who may or may not have been experienced video game players and in a non-experimental context with a sample of MMO players (Ryan, Rigby, & Przybylski, 2006). Thus, the PENS is yet to be validated with a broader (non-MMO), experienced game player sample.

## **Immersive Experience Questionnaire (IEQ)**

One aspect of player experience that is commonly referred to as important is that of immersion (Brown and Cairns, 2004). Immersion is the experience of being “in the game”, that is, being emotionally and cognitively invested in the activity of playing. There have been several approaches to studying the immersive experience had by gamers (Brockmyer et al., 2009; Qin et al., 2009; Jennett et al., 2008; IJsselsteijn et al., 2007).

In Jennett, et al, (2008) immersion is defined as “a total experience where the player fully concentrates on the game”. The five identified components of game immersion are: Cognitive involvement, barriers are effort and attention (To what extent did you feel focused on the game?), Emotional involvement (To what extent were you interested in seeing how the game’s events would progress?), Challenge (Were there any times during the game in which you just wanted to give up?) and Control (At any point did you find yourself become so involved that you were unaware you were even using controls?)

For game immersion, Jennett et al. (2008) compared the IEQ with a single score measure (rate how immersed you feel from 1-10) and found a good correlation. The IEQ is a widely used questionnaire in determining the levels of immersion experienced by players (Nacke & Lindley, 2010, Connolly et al., 2012). It has been tested much more empirically across a far-reaching array of different scenarios and game types. Similarly to the GEQ, it uses five-point Likert scale questions for measuring player experience, but is specifically focused on the notion of immersion when playing games.

## Industry questionnaire

An additional subjective rating questionnaire was tested alongside the above to test the quality of existing questionnaire used in industry. This questionnaire is used in playtests run by Sony Computer Entertainment Europe. The quantitative questions are focused on a different dimensions of PX. One question asked about overall enjoyment of the PX, while another question asked the player to give a numerical rating of the game. Three of the questions asked about key game design areas which influences sensory perception - graphics, sound effects, and music. The product evaluation side of game quality evaluation is also covered in this questionnaire. A NET promoter scale item is included, paired with a descriptive question by a prompting the players to describe the game to a friend or colleague with one word. By capturing quantitative and qualitative data in the same questionnaire, the Industry questionnaire attempts to capture a comprehensive view of player's opinion on a pre-released game. Although the questionnaires used on a regular basis during game testing, it had not been validated.

## Research questions

Overall this literature review reveals that Metacritics scores and questionnaire scores are reflecting different dimensions of player experience. While there is existing research done on the relation between the two, it cannot account for the variability of player experience in an individual player across games of different quality.

To investigate into how well questionnaires can segregate games of different quality, and how game experiences measured by various questionnaires relate to game quality (as defined by Metacritic scores), the following exploratory research questions were formulated:

1. Do subjective questionnaires (IEQ, GEQ, PENS, Industry) show differences between games of different quality?
2. Is there a relationship between subjective questionnaire scores (IEQ, GEQ, PENS, Industry) and Metacritic scores?

The answers can help tighten the gap between academic and industry practises and check which components of those questionnaires are more effective in predicting a game's quality.



# Chapter 3. Method

## Participants

Thirty-nine participants (37 male, 2 female), aged between 18 and 35 years ( $M = 25$ ,  $SD = 5.21$ ) were recruited through an opportunity sample. Participants were recruited and compensated at Sony Computer Entertainment Europe using game testing standard procedures and rates, to ensure the experimental design and conditions were as close as possible to industry common practice.

Recruitment criteria included high level of interest in First Person Shooter (FPS) games. We specifically asked for participants that had played and finished several games in the genre, to ensure a minimum of proficiency and genre knowledge. One of the recruitment criteria was that participants did not play or own any of the 5 games used in the study as an attempt to avoid the selection bias, to control the amount of experience with these games across participants.

This also brought the experiment setting closer to a traditional user test prior to a game release or during game development (Lewis-Beck et al., 2003).

## Experimental Design

A within subject design was used to explore the player experience for games with different perceived quality by game critics. All participants played all five games in this study and the quality of the game was registered by metacritic scores (over 85, 84-75, 74-65, 64-55). The game experience was measured by self-reported questionnaires. To the extent possible, the methods employed in this lab study controlled and limited the potential effects of environmental variables in the test that do not relate to the game.

## Materials and Apparatus

### FPS game selection

Five First Person Shooter (FPS) games were selected, which fell into four score buckets, which represents four different levels of quality of received critic scores. (See table 1). Games were chosen from this genre because it was among one of the top selling game genres (Statista, 2015), and that it had a larger range of games with various quality available for the game console chosen for this study. The scores were obtained from metacritic.com, a popular website that takes reviews from many

sources and normalizes the rating from each into a 100-point scale. Games which are on the top 50 best sold PS3 were excluded to avoid a selection bias in questionnaire ratings. It was assumed that the rest of the games were less well-known to mainstream players based on the sales figures. One game was chosen for each metacritic score bucket. The choice of having two games in the highest score bucket was for the interest of understanding the relationship between component scores in questionnaires and high metacritic scores. A brief description of each game, together with its screenshot, is listed in appendix 9.

Table 1. Selected games and metacritic scores

Game	Metacritic Score	Score bucket
Borderlands 2	91	over 85
Farcry 3	90	over 85
Metro Last Light	80	84-75
Homefront	70	74-65
Medal of Honor: Warfighter	55	<64

## Questionnaires

All questionnaires were loaded and answered on the tablet next to the console, and the following measures were utilised:

1. Initial questionnaire: A demographic questionnaire asked participants their age, gender and interests in different game genres, see appendix 1.
2. Pre-game questionnaire: It asked participants if they had heard of the tested game title. If they had heard of the game title, they were asked to rate the game from what they knew about this game from memory, see appendix 2.
3. Post-game questionnaire: A set of questionnaires consisted of the below:
  - 3.1. Industry questionnaire consisted of 10 questions, with 6 11-point Likert scale items and 4 text-based question, see appendix 3.
  - 3.2. Immersive Experience Questionnaire (Jennett et al., 2008) consisted of 31 7-point Likert scale items, see appendix 4.
  - 3.3. Game Experience Questionnaire (IJsselsteijn et al., 2007) consisted of 42 5-point Likert scale items, see appendix 5.
  - 3.4. Player Experience of Need Satisfaction Questionnaire (Ryan et al., 2006) consisted of 18 7-point Likert scale items, see appendix 6.

4. Final questionnaire: Participants filled this in after playing all five games, which ranked the games in order of components of each game question sets, see appendix 7.

## Representative Gameplay session selection

Since each of these games accounted for more than 10 hours of total gameplay time (GameLengths, n.d.), it was impractical for participants to experience it in its entirety before they filled in the post-game questionnaires. , it was impractical for participants to experience it in its entirety before they filled in the post-game questionnaires. After time constraints and compensation resources were considered, 60 minutes of most representative gameplay for each game was selected to provide participants a good account of the gameplay experience for each of the game, even though they would not be able to experience the game in full. The selection was done by excluding the below described inappropriate sessions and including suitable levels of the game, together with modifications from pilot study. Since the genre being tested was First Person Shooter, the levels where the player wasn't required to perform core FPS gameplay (e.g. moving and foot and shooting) were considered as not appropriate for this study. This is because these levels could not provide the representative gameplay of a FPS game and hence excluded in the gameplay session. The gameplay sessions were chosen one chapter after the tutorial or initial chapter so participants could skip the basics and experience the typical gameplay of each game.

## Lab set-up

The 39 participants were tested in groups of 10, where participants were seated in independent and identical gaming pods. Each pod was equipped with a PlayStation 3 console, Playstation 3's standard Dualshock 3 controller and a tablet for answering questionnaires (see appendix 8). There was a game instruction sheet in each pod which contained a background story of the game and details of the gameplay session they were about to play (see appendix 9). Each participant was given an instruction sheet to give them an idea of the gameplay schedule (see appendix 10). The questionnaires were administered on the tablet next to the console.

## Procedures

All participants were treated in the same way from the time they were contacted about participating to the time they leave the lab, and personnel (researcher and her supervisor) conducting playtests were thoroughly trained to ensure that they treated participants consistently. Participants were greeted and

welcomed into the testing area, followed by an introduction to the study and the signing of the consent form (see appendix 11). Participants were then led to their respective gaming pods. Initial questionnaire was filled in by participant at the beginning of the study.

Before the start of every gameplay experience participants were asked to fill in the pre-game questionnaire. Participants then started the one hour of representative gameplay session on the Playstation 3 in their respective pod. For all games participants were instructed to play the game in their most natural setting, “as if they’re at home” and just reached the level they were about to play. They were instructed to play individually (without interacting with others) and answer the same question about the game as they play. Immediately after the one-hour gameplay participants were asked to fill in the post-game questionnaire, which follows the practice in typical playtests (Davis, 2005). This process repeated for five times, and the order of the games were counterbalanced using Latin Square across participants to control for order effects. (See appendix 12) The post-game questionnaire was randomised so that each questionnaire in the post-game questionnaire was presented in a random order, and its items were also randomised, to counterbalance the order effect. After playing all five games, participants filled in the final questionnaire Then they were debriefed on the goals of the study and received monetary compensation.

## Pilot study

Before the recruitment of a larger participant sample for the main experiment, a pilot study was carried out with four participants, whom were regular gamers of FPS. It aimed to evaluate the fluency of the main experiment, and to check the fatigue level of participants after filling in the same set of questionnaires several times a day. The pilot study consisted of a short gameplay session for six games and questionnaire filling, followed by a focus group where players were invited to share their experience on the six games. They were asked to describe their subjective fatigue level verbally with short sentences and adjectives after playing six games on the same day, and on the length of the questionnaires to answer after each gameplay.

Adjustments were made based on questionnaire scores and focus group discussion results from the pilot study. It was found that participants rated Aliens colonial marines based on their expectations from the popular movie franchise, hence this game was removed from the main experiment to avoid experimental bias. The representative gameplay session selection was modified according to the questionnaire scores and focus group discussion results, by increasing the difficulty level for Medal of

Honor, limiting the number of missions to be done in Borderlands 2 and Farcry 3, and selecting a later chapter of Metro Last Light.

## Chapter 4. Results

After collecting the questionnaire data, the statistics were calculated using SPSS 22.

### Pre-game questionnaire

Below are the pre-game questionnaire results (see table 1). Most of the participants had heard of Farcry3 (94.9%). Metro Last Light is the game that least of the participants had heard of (41%). Based on what participants knew about this game, Farcry was rated the highest (8.22), and Homefront the least (7.03).

	Borderlands 2	Farcry 3	Metro Last Light	Homefront	Medal of Honor: War Fighter
Have you heard of this game before?					
Yes (In percentage)	89.7	94.9	41	79.5	64.1
Based on what you know about this game, what would you rate this game?					
Mean	7.63	8.22	7.06	7.03	7.48
Std. Deviation	2.03	1.685	2.235	1.472	1.388

Table2. Mean and SD of Pre-game Questionnaire

### Differences between games of different quality

To answer the research question, “Do subjective questionnaires (IEQ, GEQ, PENS, Industry) show differences between games of different quality?”, the questionnaire scores were computed into total scores or component scores according to the original studies from the questionnaire developers (IJsselsteijn et al., 2007; Jennett et al., 2008; Ryan et al., 2006). The highest and lowest scoring game in each component is identified.

To find out whether subjective questionnaires (IEQ, GEQ, PENS, Industry) show differences between games of different quality, a repeated ANOVA was carried out. Repeated Measures ANOVA detects any overall differences between related means.

The Mauchly's Test of Sphericity was carried out to determine whether assumption of sphericity is violated for each component, and Greenhouse-Geisser Correction was applied on components which violated the assumption.

Below presents the results by questionnaire.

## GEQ

The means, standard deviations for the components of GEQ are displayed in Table 2 , and the highest and lowest scoring game in each component is displayed in Table 3. Mixed results is observed in GEQ, where all games scored highest in different components (except Homefront), with Farcry 3 dominated the highest scoring game in several components (Sensory and Imaginative Immersion: 3.82, Challenge: 3.23, Positive affect: 3.39) . In GEQ. Homefront did not get the highest score in any components and got the lowest score in the majority of components.

There is a significant effect of games with different Metacritic scores on the following components: Competence ( $F(4, 152) = 2.689, p < .0005$ ), Sensory and Imaginative Immersion ( $F(3.033, 136.74) = 7.584, p < .0005$ ) and Challenge ( $F(4, 152) = 4.915, p < .0005$ ) in GEQ.

Game	Borderlands 2		Farcry 3		Metro Last Light		Homefront		Medal of Honor: War Fighter		
	M	SD	M	SD	M	SD	M	SD	M	SD	
Metacritics Professional Score	91.00		90.00		80.00		70.00		55.00		
Means and Standard Deviation	M	SD	M	SD	M	SD	M	SD	M	SD	
GEQ	Competence	3.05	0.88	3.23	0.99	3.07	0.82	3.27	0.94	3.54	1.00
	Sensory and Imaginative Immersion	3.03	1.05	3.82	0.77	3.12	1.02	2.81	0.99	3.01	1.06
	Flow	3.30	1.10	3.39	1.15	3.24	1.13	3.08	1.07	3.41	1.19
	Tension	1.97	0.90	1.93	0.84	2.02	0.86	1.80	1.00	1.97	0.88
	Challenge	3.01	0.85	3.23	0.79	2.89	1.03	2.59	0.88	3.01	0.87
	Negative affect	2.11	0.83	1.77	0.82	1.91	0.94	1.80	0.90	1.98	0.88
	Positive affect	3.12	1.08	3.39	0.93	3.08	1.01	3.27	0.77	3.22	0.87

Table3. The means and standard deviation of each component of GEQ

		Highest	Lowest
GEQ	Competence	Medal of Honor: War Fighter	Borderlands 2
	Sensory and Imaginative Immersion	Farcry 3	Homefront
	Flow	Medal of Honor: War Fighter	Homefront

	Tension	MetroLast Light	Homefront
	Challenge	Farcry 3	Homefront
	Negative affect	Borderlands 2	Farcry 3
	Positive affect	Farcry 3	Metro Last Light

Table4. Highest and lowest scoring game in each component of GEQ

		Mauchly's Test of Sphericity Sig.	Sphericity or Correction	F	Sig.
GEQ	Competence	0.365	Sphericity Assumed	2.689	0.033
	Sensory and Imaginative Immersion	0.015	Greenhouse-Geisser	7.584	0
	Flow	0.34	Sphericity Assumed	0.859	0.49
	Tension	0.102	Sphericity Assumed	0.426	0.79
	Challenge	0.222	Sphericity Assumed	4.195	0.003
	Negative affect	0.885	Sphericity Assumed	1.168	0.327
	Positive affect	0.226	Sphericity Assumed	0.798	0.528

Table 5. Repeated ANOVA results for GEQ

## IEQ

The means, standard deviations for the components of IEQ are displayed in Table 4 , and the highest and lowest scoring game in each component is displayed in Table 5. Farcry 3 scored highest over other games in IEQ questionnaire. Homefront and Borderlands 2 each scored the lowest in two of the components.

There is a significant effect of games with different Metacritic scores in Control ( $F(4, 152) = 5.882, p < .0005$ ) in IEQ.

	Game	Borderlands 2		Farcry 3		Metro Last Light		Homefront		Medal of Honor: War Fighter	
		Metacritics Professional Score		90.00		80.00		70.00		55.00	
	Means and Standard Deviation	M	SD	M	SD	M	SD	M	SD	M	SD
IEQ	Total	135.33	36.42	147.26	28.79	140.00	33.36	133.51	29.88	142.69	36.56
	Cognitive Involvement	49.79	13.01	51.97	10.16	50.46	11.67	49.97	10.30	51.72	11.93
	Real World Dissociation	24.69	7.46	26.13	7.46	24.79	7.82	23.36	6.33	25.95	8.06

	Emotional Involvement	48.18	16.28	54.67	13.49	52.77	15.88	50.97	13.96	52.67	15.56
	Challenge	19.90	3.35	20.33	3.47	20.33	3.86	19.92	4.11	19.64	3.58
	Control	32.26	8.11	35.92	7.84	33.31	8.76	28.72	9.29	32.72	8.16

Table6. The means and standard deviation of each component of IEQ

		Highest	Lowest
IEQ	Total	Farcry 3	Homefront
	Cognitive Involvement	Farcry 3	Borderlands 2
	Real World Dissociation	Farcry 3	Homefront
	Emotional Involvement	Farcry 3	Borderlands 2
	Challenge	Farcry 3, Metro Last Light	Medal of Honor: War Fighter
	Control	Farcry 3	Homefront

Table7. Highest and lowest scoring game in each component of IEQ

		Mauchly's Test of Sphericity Sig.	Sphericity or Correction	F	Sig.
IEQ	Total	0.238	Sphericity Assumed	1.37	0.25
	Cognitive Involvement	0.055	Sphericity Assumed	0.398	0.81
	Real World Dissociation	0.017	Greenhouse-Geisser	1.406	0.242
	Emotional Involvement	0.382	Sphericity Assumed	1.232	0.3
	Challenge	0.996	Sphericity Assumed	0.483	0.748
	Control	0.056	Sphericity Assumed	5.882	0

Table 8. Repeated ANOVA results for IEQ

## PENS

The means, standard deviations for the components of PENS are displayed in Table 9 , and the highest and lowest scoring game in each component is displayed in Table 10.

Mixed results is observed in PENS, where all games scored highest in different components, with Farcry 3 is the highest scoring game in two of the four components.

There is a significant effect of games with different Metacritic scores on the following components: In-Game Autonomy ( $F(3.221, 122.39) = 26.056, p < .0005$ ), Presence ( $F(3.192, 121.31) = 2.824, p < .0005$ ) and Intuitive Control ( $F(4, 152) = 5.055, p < .0005$ ).

	Game	Borderlands 2		Farcry 3		MetroLast Light		Homefront		Medal of Honor: War Fighter	
		M	SD	M	SD	M	SD	M	SD	M	SD
	Metacritics Professional Score	91.00		90.00		80.00		70.00		55.00	
PENS	In-Game Autonomy	4.38	1.65	5.91	1.15	3.74	1.55	3.32	1.52	3.30	1.63
	In-Game Competence	4.78	1.58	4.85	1.39	4.82	1.28	5.32	1.42	5.21	1.37
	Physical/Emotional/Narrative Presence Scale	3.05	1.20	3.80	1.20	3.50	1.43	3.40	1.25	3.34	1.55
	Intuitive Control	5.56	1.35	5.50	1.50	4.78	1.32	4.93	1.42	4.79	1.42

Table11. The means and standard deviation of each component of PENS

		Highest	Lowest
PENS	In-Game Autonomy	Farcry 3	Medal of Honor: War Fighter
	In-Game Competence	Homefront	Borderlands 2
	Physical/Emotional/Narrative Presence Scale	Farcry 3	Borderlands 2
	Intuitive Control	Borderlands 2	Metro Last Light

Table12. Highest and lowest scoring game in each component of PENS

		Mauchly's Test of Sphericity Sig.	Sphericity or Correction	F	Sig.
PENS	In-Game Autonomy	0.011	Greenhouse-Geisser	26.056	0
	In-Game Competence	0.054	Sphericity Assumed	1.623	0.171
	Presence	0.036	Greenhouse-Geisser	2.824	0.038
	Intuitive Control	0.06	Sphericity Assumed	5.055	0.001

Table 13. Repeated ANOVA results for PENS

## Industry Questionnaire

The means, standard deviations for the components of Industry Questionnaire are displayed in Table 11, and the highest and lowest scoring game in each component is displayed in Table 12. Farcry 3 scored highest over other games in all components. Borderlands 2 scored the lowest in three out of six of the components. There is a significant effect of games with different Metacritic scores on Graphics ( $F(4, 152) = 7.522, p < .0005$ ) in Industry questionnaire.

	Game	Borderlands 2		Farcry 3		MetroLast Light		Homefront		Medal of Honor: War Fighter	
		M	SD	M	SD	M	SD	M	SD	M	SD
	Metacritics Professional Score	91.00		90.00		80.00		70.00		55.00	
	Means and Standard Deviation	M	SD	M	SD	M	SD	M	SD	M	SD
Industry	How would you RATE the game?	6.38	2.46	7.54	1.75	6.59	2.48	6.56	1.67	6.79	2.15
	How would you rate the GRAPHICS in the game?	5.87	2.65	7.59	1.83	6.67	2.36	5.23	2.36	6.82	1.96
	How would you rate the SOUND EFFECTS in the game?	6.64	2.08	7.28	1.50	6.82	2.69	6.54	1.88	6.79	2.55
	How would you rate the MUSIC of the game?	6.31	2.58	7.05	1.73	6.26	2.49	6.10	1.96	6.03	2.38
	Overall, I ENJOYED playing the game	6.49	2.73	7.79	1.72	6.92	2.62	6.87	1.70	7.08	2.26
	NET promoter score	5.85	3.13	7.23	2.37	6.38	2.91	6.15	2.43	6.62	2.72

Table14. The means and standard deviation of each component of Industry Questionnaire

		Highest	Lowest
Industry	How would you RATE the game?	Farcry 3	Borderlands 2
	How would you rate the GRAPHICS in the game?	Farcry 3	Homefront
	How would you rate the SOUND EFFECTS in the game?	Farcry 3	Homefront
	How would you rate the MUSIC of the game?	Farcry 3	Medal of Honor: War Fighter
	Overall, I ENJOYED playing the game	Farcry 3	Borderlands 2
	NET promotor score	Farcry 3	Borderlands 2

Table15. Highest and lowest scoring game in each component of Industry Questionnaire

		Mauchly's Test of Sphericity Sig.	Sphericity or Correction	F	Sig.
User Rating	Overall Rating	0.027	Greenhouse-Geisser	2.08	0.10
	Graphics	0.516	Sphericity Assumed	7.522	0
	Sound Effect	0.21	Sphericity Assumed	0.841	0.501
	Music	0.295	Sphericity Assumed	1.964	0.103

	Overall Enjoyment	0.001	Greenhouse-Geisser	2.083	0.086
	NET promoter score	0.162	Sphericity Assumed	1.604	0.176

Table16. Repeated ANOVA results for Industry Questionnaire

## Post Hoc Bonferroni Correction

Post hoc tests using the Bonferroni correction revealed that several scores for Farcry 3 across questionnaires and components were significantly different from most of the other games, for instance, Control in IEQ (M=4, SD = 1.29) between medal of honor and homefront. None of the questionnaires showed significant differences between all games of different quality, as not all scores of the five games were found to be significantly different from the rest of the games. (See table 15) It is interesting to see that component scores for Farcry 3 and Borderlands 2 were significantly different across several questionnaire and components, despite having very close Metacritic scores. The differences of GEQ Competence component scores between the five games were not significant. For a detailed comparison, please see the table 17 below, and for a full table of post hoc Bonferroni Test please see Appendix 13.

Questionnaire	Measure	(I) Games	(J) Games	Sig.
IEQ	Control	Medal of Honor	Homefront	0.036
		Farcry 3	Homefront	0.000
GEQ	Sensory and Imaginative Immersion	Farcry 3	Medal of Honor	0.001
			Homefront	0.000
			Metro Last Light	0.004
			Borderlands 2	0.001
	Challenge	Farcry 3	Homefront	0.004
PENS	In-Game Autonomy	Farcry 3	Medal of Honor	0.000
			Homefront	0.000
			Metro Last Light	0.000
			Borderlands 2	0.000
	Borderlands 2	Medal of Honor	0.03	
		Homefront	0.013	
	Presence	Farcry 3	Borderlands 2	0.021

	Intuitive Control	Medal of Honor	Metro Last Light	0.043
			Borderlands 2	0.03
Industry	Graphics	Farcry 3	Homefront	0.032
			Metro Last Light	0.021
			Borderlands 2	0.000

Table 17. Significant components in post hoc Bonferroni Test

## Final Questionnaire

	Borderlands 2	Far Cry 3	Metro last light	Homefront	Medal of honor: warfighter
Order the five games from your most favourite to least favourite.					
Mean	3.1282	2.641	2.8205	3.2564	3.1538
Std. Deviation	1.52487	1.42325	1.41183	1.27151	1.4242
Order the five games from most immersive to least immersive.					
Mean	3.3846	2.4359	2.8205	3.4359	2.9231
Std. Deviation	1.47996	1.33367	1.58731	1.14236	1.32555
Order the five games from best game experience to worst game experience.					
Mean	3.4359	2.3333	3	3.4103	2.8205
Std. Deviation	1.5525	1.34425	1.37649	1.22942	1.33519
Order the five games from the one you most likely to play again to the least likely to play again...					
Mean	3.5128	2.5641	2.9231	3.0513	2.9487
Std. Deviation	1.43034	1.48315	1.49358	1.16864	1.39451
Order the five games from the one you most likely to recommend to your friends and colleagues to the least					
Mean	3.4103	2.4359	3	3.2051	2.9487
Std. Deviation	1.5512	1.4653	1.39548	1.03057	1.46806
Order the five games from the most fun to the least fun.					
Mean	3.0513	2.5641	3.3077	3.0513	3.0256
Std. Deviation	1.57194	1.39161	1.36022	1.25549	1.45976
<b>Overall Rank</b>	3.320516667	2.495716667	2.978633333	3.23505	2.970066667

The final questionnaire data was analysed. All rank-scores were averaged and a final rank is calculated. At the end the lowest average is the first choice. (rank 1, gets coded 1... rank 5 is 5). The

questions have similar ranks for games. The question, “Order the five games from most immersive to least immersive.” and “Order the five games from best game experience to worst game experience.” has a similar ranking of games. However, the question “Order the five games from the one you most likely to recommend to your friends and colleagues to the least” is different, that Borderlands 2 is ranked as lowest while Farcry 3 scored highest in the NET promoter item. This gives an overall rank (by averaging the means from the ranks) from the final questionnaire: Farcry 3 in the highest rank, followed by Metro Last Light, Medal of honor: warfighter, Homefront, and Borderlands 2.

## Relationship between Metacritic Scores and component scores

Pearson correlation analysis was carried out to answer the research question “Can the scores given to these components be used to predict Metacritic score?” Pearson correlation coefficient is a measure of the strength of a linear association between two variables, which indicates how well the data points fit the line of best fit. This was used to check whether there was a linear relationship between the Metacritic Scores and the questionnaire component scores.

Analysis with Pearson Correlation (Table 3) reveals significant positive correlations between Metacritic Scores and In-Game Autonomy in PENS ( $r=.41, p<.05$ ) and immersion in GEQ ( $r=.18, p<.05$ ) and Metacritic scores. There were significant weak negative correlations between Intuitive Control in PENS ( $r=-.21, p<.05$ ) and Competence in GEQ ( $r=-.16, p<.05$ ) and Metacritic scores. None of the component scores in IEQ or Industry Questionnaire had a significant Pearson correlation with the Metacritic scores.

Based on the above analysis, significant linear relationships between Metacritic Scores and some of the components were observed.

Table 5. Pearson Correlation with Metacritic Scores

GEQ		IEQ	
Competence	0.084	Total	0.074
Sensory and Imaginative Immersion	.172*	Cognitive Involvement	0.048
Flow	0.003	Real World Dissociation	0.027
Tension	-0.049	Emotional Involvement	0.105
Challenge	0.021	Challenge	0.018
Negative affect	-0.128	Control	0.045
Positive affect	0.113		

PENS		Industry	
In-Game Autonomy	.209**	Overall Rating	.151*
In-Game Competence	0.063	Graphics	0.12
Prescence	.149*	Sound Effect	0.065
Intuitive Control	0.099	Music	0.084
		Overall Enjoyment	.160*
		NET promoter score	0.131

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

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

## Chapter 5. Discussion

### **Do subjective questionnaires (IEQ, GEQ, PENS, Industry) show differences between games of different quality?**

Across all four questionnaires, differences in game component scores were found with Control in IEQ, Intuitive Controls and Autonomy in PENS, and Challenge and Competence in GEQ, which are related to player's ability to control and navigate within the game. This means that these components could segregate the games by scores. These differences across games may reflect that these aspects of the game are a key area for professional critics to separate good games from poor ones. This can be explained by the genre of the study, where autonomy and sense of control are key components of First Person Shooter videogames.

There are differences in game component scores in Sensory and Imaginative Immersion in GEQ and Physical/Emotional/Narrative Presence in PENS, which refer to a more emotional, subjective aspects of player experience.

Three out of four components of the PENS, five out of seven components of the GEQ and one component of the IEQ showed significant differences between the component scores of the five games with different Metacritic professional scores.

The higher number of PENS components having differences in component scores compared to other questionnaires could reflect the relative importance of the satisfaction of self-determination theory related needs in creating a rewarding player experience. At the very least, the constructs measured by the PENS measure seem to be of high importance for people (reviewers and users) that rank videogames on Metacritic. It may also be that the PENS is a relatively stronger measure in terms of psychometric properties, having been formally validated and refined in various settings.

Despite both being in the top-scored category based on Metacritic Scores, Borderlands and Far Cry are have very different component scores. These two games are similar games in terms of open world/sandbox nature, and it is surprising to see such results. One possible explanation could be due to the sub-genre of the games. Borderlands is more quest based and RPG-like, while Far Cry has a lot of different gameplay elements which features on autonomy and control, which are valued highly by FPS players.

In the final questionnaire, the results provided a good summary of their experience in ranks - players in general favoured Farcry 3 most. In GEQ, Homefront scored the lowest amongst all components, while in other questionnaires the lowest scored games were close to evenly distributed amongst Medal of Honor, Homefront, Borderlands 2 and Metro Last Light. It might mean that the questionnaires are in general perform better, with results align with each other, in identifying games that players prefer, and on the other hand perform worse in identifying games that players least prefer.

## **Is there a relationship between subjective questionnaire scores (IEQ, GEQ, PENS, Industry) and Metacritic scores?**

By comparing Metacritic Professional Scores with the components of player experience measured in the current study, some clear differences and similarities are observed between what professionals and players are responding to.

Across all four questionnaires, significant correlations between component scores and Metacritic Professional scores were found with In-Game Autonomy and Intuitive Control in PENS, Sensory and Imaginative Immersion and Competence in GEQ. This suggests that higher Metacritic scores are indicative of higher component scores, which suggests that these components have an influence on professional critics' game review.

PENS In-Game Autonomy and GEQ Sensory and Imaginative Immersion are related to the scores given by players and the Metacritic Professional Score. It is likely that these are universally positively regarded components of player experience.

There is no relationship between the IEQ component scores and the Metacritic Professional Scores. There might be a difference between the formalized idea of immersive experience developed in academia and that of game critics. This may reflect that the aspects quantified in IEQ are not the focuses of professional critics. This pattern may reflect a lesser ability on the part of professional critics to identify how immersive a player can be in a game. Professional reviewers arguably have other focuses than immersion or overall enjoyment. Professional reviewers arguably have other focuses than immersion or overall enjoyment, such as the above mentioned In-Game Autonomy and Intuitive Control in PENS.

There is a negative correlation between two of the Industry questionnaire item scores and the Metacritic Professional Scores. This might reflect the opposite perspective of enjoyment in professional critics and players.

Overall, it can be seen that Metacritic professional scores relate to some of the player experience facets that are measured by the IEQ, PENS, GEQ and the Industry questionnaire.

This suggests there is overlap between player experience constructs being assessed these questionnaires and the game aspects to which professional critics react to. Metacritic professional scores are more strongly associated with the components of the PENS and GEQ than the IEQ or Industry questionnaire. The different results across all questionnaires might verify the assumption that they all interpret player experience and flow theory differently.

To conclude, there is a relationship between some of the components in the subjective questionnaires (IEQ, GEQ, PENS, Industry) and Metacritic scores. This is partially in line with the results in Johnson et al. (2014), with a lower number of significant effects. In the study by Johnson et al., all components of PENS was strongly correlated with Metacritic Professional Scores, where only half of the components had correlations with Metacritic Professional scores in this study.

Based on the results to the two research questions, the Metacritic scores cannot be seen as having predictive power on player experience, nor the player experience components defined by academic researchers matched how the professional game critics evaluate a game. The dichotomy between academic research and business perspectives on game quality evaluation is highlighted in this study, but further research would need to be done in order to investigate into the roots of this dichotomy.

## Strengths in current research

A strong point of this study is on the repeated measure design. To compare the results of the four questionnaires, the same sample of people completed all four player experience measures. By assuming participants rated all five games with the same personal criteria, the impact of inter-individual differences is reduced compared to a non-repeated measure design. Another strong point is that the study was carried out in a lab environment and well controlled.

This study is the first attempt in comparing the IEQ, GEQ, PENS and a questionnaire developed within the industry, and the first one in comparing IEQ scores with Metacritic scores. This opened up future possibilities in adapting a similar methodology in researching into standardised tools in subjective player experience evaluation and bridging the gap between academic research and business perspectives.

# Limitations and Confoundings

## Current restrictions in studying player experience

Researchers are still looking to generate a common ground of what game experience means. Various, overlapping concepts within player experience (PX) make it difficult to develop valid measures of central components such as game enjoyment, flow, presence, immersion etc. However currently there is not enough literature, research or tools to validate the ratings in questionnaires, to identify the mismatch between rationale behind how a user rates and the rating value given, or to compare the central PX components measured in each questionnaire.

## Limitations of questionnaires as a research methodology

However there are problems on the use of questionnaires.

### Meaning of a numerical score

While a numeric scale is intuitive in many respects and allows for relative judgments, information describing how the numeric score translates into an absolute judgment of their experiences is not known. There is discrepancy between how people rate a game and how these ratings are analysed.

The metacritic scores of the five selected games did not span the entire 100 point range.

The Metacritic Scale is far from linear - most games score around 70, very few score above 90 and almost none score under 50. However, correlation methods used looks for a linear relationship, and this might be the cause of no significant correlation was found.

### Biases when filling in questionnaires

However there is a central bias tendency with Likert scales for people to respond towards the middle of the scale perhaps to make them look less extreme. This might contribute to the inability in segregating the games using Likert-scale based questionnaire.

Moreover, there are a wealth of documentation on the biases and other difficulties subjected to questionnaires, self-reporting method, for example, social desirability bias (Crowne & Marlowe, 1960) described the tendency that participants answer in a way that will be positively recognised. As with any questionnaire participants may provide the answers that they feel they should, despite at the

beginning of an experiment they had been explained that the how they rate the quality of the game had no conflict of interest with the researchers. This might explain the results are at the higher ends of the numeric scale.

## Retrospective

Questionnaires are also retrospective in nature. When players are asked in a survey to give their views about a game, they are asked to think about their past experiences with a game, experiences that may have varied widely. Survey respondents may have finished the game, quit because it was not fun or played very little of the game. The retrospection also induces a rationalisation of the experience rather than a direct measurement of the experience. In this study participants were asked to play a specific section of the game, where they were stopped they reached the end of the pre-designated end point or when the playtime of 60 minutes was finished, this meant that some participants didn't have the exact same experience as others. For example, participants could be asked to stop at a point when they struggled to pass a tricky section, and this might lead to their play experience ending on a potentially negative note because they were failing. Some other users finished the experience and maybe finished on a end note (e.g. finish a mission by killing a boss). The recency effect may have had an impact on the way they reported their experience as the last part of the experience can have been quite different between users. It is often not clear, therefore, that survey respondents are reporting their perceptions of the same gaming experience, which makes it difficult to draw generalizable conclusions from data collected in this manner.

Question writing, either close or open questions, is difficult to yield actionable data.

the wording of the questions themselves that reduce the face validity (Adams & Cox, 2008)

Since the data is mainly quantitative (except the 3 qualitative questions in Industry Questionnaire), it does not provide in depth replies of the players' subjective experience. Lazzaro and Mellon (2005) highlighted that triangulating and correlating between data sources, whether all quantitative or mixed quantitative/qualitative, improves the validity and applicability of results.

## Experimental Confoundings

Overall, the effect sizes associated with the relationships identified in the present study are relatively small. This suggests that the factors identified by academic researchers as being key components of player experience are having only a small impact on review scores. It may also be that the influence of these factors is weakened as part of the weighting process of professional scores undertaken on the Metacritic website. For example, if specific critics treated as less influential (and therefore

apportioned relatively less weight as part of the overall professional Metacritic score) are more influenced by these factors than critics considered more influential, the strength of the relationship would be partially hidden. This might be the case if more influential critics are, for example, more jaded, more cynical or more subjectively critical than less influential or less experienced critics.

Another reason for relatively small number of significant effects between player experience and metacritic scores and its weak game quality segregation can be explained by noises or biases. As some of the participants had heard of the games before (referring to pre-test questionnaire results), a post-play cognitive rationalization of the experience with the content of the review might have positively biased the questionnaire results. Another reason is that the players only had one hour to play each of the game, which might not be enough to establish a good enough judgement for each of the games. All the games were played in one day and it might have caused fatigue, especially they were asked to fill in a questionnaire after each gameplay, where similar wording are used in questionnaires. Despite the best effort in recruiting a sample that is representative of the FPS gamer population, the study is restricted by the sample size, subjected to noise and confounding variables, and therefore difficult to enable meaningful statistical analyses and comparisons.

If players were allowed to spend longer playing a game, then it might have influenced component scores, namely those related to control and immersive experience, as they become familiar with controls even if they are not intuitive. Since *Borderlands 2* had more complicated loot system and complicated inventory (Metacritics, 2015), spending longer in the game environment might help players become more intuitive to the controls and enjoy the autonomy and freedom in *Borderlands 2* as a game with sandbox genre elements.

The limited ability of the questionnaire to segregate games with different Metacritic score can be due to the fact that full range of values on the measurement scales was not included in the current study. The Metacritic Score of tested games only ranged from 55 to 91, and the limited range may have caused an underestimate the true relationships from the reported results. Future research should extend to games that are relatively less popular.

## Future research and implication

Increasing the collaboration between academic research and commercial playtest, or game quality validation across-business functions within the same company, can be useful in bridging the existing gap between player experience approach and product quality approach in assessing game quality.

To validate existing questionnaires, further research is needed in testing questionnaire sets on games in different genres in all ranges of game quality (whether determined by Metacritics such as this study, or other indicators), with a larger pool of participants. To further investigate the relationship between

component scores and Metacritic Professional Scores, researchers can carry out interviews with professional game critics (who contributes to Metacritic Professional Scores) and capture their view of player experience, and contrast it with consumers of the games.

Developing a set of consistent and standardised tool in evaluating game quality is crucial. Having a widely adopted standard test game (or a set of games with various game quality) can be useful as a benchmarking for game researchers, similar to the use of Lenna image in the field of image processing since 1973. (Gonzalez, Woods & Eddins, 2004)

The components of player experience might have different weights or level of importance in determining player experience in different genres. For example, control might be more important in a FPS game while immersion and real word disassociation might play a more important role in a action-adventure game. Since this research is only done with one game genre, First Person Shooter, further research can be done with the same four sets of questionnaires but in other genres to look at the predictive power of Metacritic scores in other genres and eventually across all genres.

When using questionnaire scores, Lazzaro and Mellon (2005) highlighted that triangulating and correlating between data sources, whether all quantitative or mixed quantitative/qualitative, can improve the validity and applicability of results. Alongside the above mentioned research directions, attention can be placed on developing approaches to analyse and integrate results from different data sources and translate these results into actionable solutions for game developers and designers.

## Chapter 6. Conclusion

In this thesis we have examined subjective ratings in questionnaires as measures of player experience in relation to Metacritic Professional scores, and how well these measures are able to segregate games of different Metacritic scores. This is done through the use of five games of various quality, according to Metacritic scores, and comparing the self-reported subjective ratings scores.

Our findings suggest that Metacritic scores do not reflect the full complexity of player experience. Across all four questionnaires, significant correlations with Metacritic Professional scores were found only with some of the components, which were In-Game Autonomy and Intuitive Control in PENS, Sensory and Imaginative Immersion and Competence in GEQ. Significant differences in game component scores among different games were found only in components related to player's ability to control and navigate within the game, and none of the questionnaires or components could significantly segregate the games (in four different Metacritic Score buckets). This suggests these components have a limited influence on professional critics' game review, and the dichotomy between academic research and business perspectives had become more apparent through this study.

Our results are important for understanding player experience and to the game industry where reviews and user forums affect a game's commercial success. Further research, both in subjective player experience approach and product quality evaluation approach, can be done to allow more direct statistical assessment of possible relationships between metacritic scores and player experience, and the ability of quality segregation of questionnaires. It would be beneficial to replicate the study with games of different genres to compare and contrast the influence of each of these components on player experience.

# Reference

- Adams, A., & Cox, A. L. (2008). Questionnaires, in-depth interviews and focus groups. In *Research methods for human-computer interaction*.
- Albert, W., & Tullis, T. (2013). *Measuring the user experience: collecting, analyzing, and presenting usability metrics*. Newnes. (p. 108)
- Anderson, C. A., Shibuya, A., Ihori, N., Swing, E. L., Bushman, B. J., Sakamoto, A., & Saleem, M. (2010). Violent video game effects on aggression, empathy, and prosocial behavior in eastern and western countries: a meta-analytic review. *Psychological bulletin*, 136(2), 151.
- Bangor, A., Kortum, P., & Miller, J. (2009). Determining what individual SUS scores mean: Adding an adjective rating scale. *Journal of usability studies*, 4(3), 114-123.
- Bernbeck, S (2015) What drives a review score?  
<http://www.gamesindustry.biz/articles/2015-02-09-what-drives-a-review-score>
- Bernhaupt, R. (Ed.) (2010) *Evaluating User Experience in Games: Concepts and Methods*. London: Springer
- Berry, M. J., & Linoff, G. S. (2004). *Data mining techniques: for marketing, sales, and customer relationship management*. John Wiley & Sons.
- 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.
- Brown, E., & Cairns, P. (2004, April). A grounded investigation of game immersion. In CHI'04 extended abstracts on Human factors in computing systems (pp. 1297-1300). ACM.
- Brown, E., Cairns, P., 2004. A grounded investigation of immersion in games. in: ACM CHI 2004, pp. 1297–1300.
- Cacioppo, J. T., Tassinary, L. G., & Berntson, G. (Eds.). (2007). *Handbook of psychophysiology*. Cambridge University Press.
- 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.
- Cox, J. (2014). What makes a blockbuster video game? an empirical analysis of US sales data. *Managerial and Decision Economics*, 35(3), 189-198.
- Cox, J., What Makes a Blockbuster Video Game? An Empirical Analysis of US Sales Data. *Managerial and Decision Economics* 35, 3 (2013), 189-198.
- Crowne, D. P., & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology*, 24, 349-354.
- Davis, J. P., Steury, K., & Pagulayan, R. (2005). A survey method for assessing perceptions of a game: The consumer playtest in game design. *Game Studies*, 5(1).

Davis, J. P., Steury, K., & Pagulayan, R. (2005). A survey method for assessing perceptions of a game: The consumer playtest in game design. *Game Studies*, 5(1).

DFC intelligence forecasts global video game software industry to reach \$100b in 2019. (2015, May 4). Retrieved August 29, 2015, from <http://www.dfcint.com/wp/dfc-intelligence-forecasts-global-video-game-software-industry-to-reach-100b-in-2019/>

Everiss, B. Metacritic has changed the games industry. 2008, Available from: <http://www.bruceongames.com/2008/06/04/metacritic-has-changed-the-games-industry/>.

Eysenck SBG, Eysenck HJ, Barrett P (1985) A revised version of the psychoticism scale. *Pers Individ Differ* 6(1):21–29

GameLengths (n.d.). Average playtime <http://www.gamelengths.com/games/playtimes>

Garland, R. (1991). The mid-point on a rating scale: Is it desirable. *Marketing bulletin*, 2(1), 66-70.

Gonzalez, R. C., Woods, R. E., & Eddins, S. L. (2004). *Digital image processing using MATLAB*. Pearson Education India.

Hassenzahl, M., & Tractinsky, N. (2006). User experience-a research agenda. *Behaviour & information technology*, 25(2), 91-97.

IJsselsteijn, W., de Kort, Y., Poels, K., Jurgelionis, A., and Bellotti, F. Characterising and measuring user experiences in digital games. In *Advances in Computer Entertainment Technology, ACE 2007 (2007)*. Workshop 'Methods for evaluating games'.

IJsselsteijn, W., de Kort, Y., Poels, K., Jurgelionis, A., Bellotti, F., 2007. Characterising and measuring user experiences in digital games. in: *Advances in Computer Entertainment Technology, ACE 2007*. Workshop 'Methods for evaluating games'.

Jennett, C., Cox, A., Cairns, P., Dhoparee, S., Epps, A., Tijs, T., 2008. Measuring and defining the experience of immersion in games. *International Journal of Human-Computer Studies* 66 (9), 641–666.

Johnson, D., & Gardner, J. (2010, November). Personality, motivation and video games. In *Proceedings of the 22nd Conference of the Computer-Human Interaction Special Interest Group of Australia on Computer-Human Interaction*(pp. 276-279). ACM.

Johnson, D., Watling, C., Gardner, J., & Nacke, L. E. (2014, October). The edge of glory: The relationship between metacritic scores and player experience. In *Proceedings of the first ACM SIGCHI annual symposium on Computer-human interaction in play* (pp. 141-150). ACM.

Jones, C. M., Scholes, L., Johnson, D., Katsikitis, M., & Carras, M. C. (2014). Gaming well: links between videogames and flourishing mental health. *Frontiers in psychology*, 5.

Keiningham, T. L., Cooil, B., Andreassen, T. W., & Aksoy, L. (2007). A longitudinal examination of net promoter and firm revenue growth. *Journal of Marketing*, 71(3), 39-51.

Lazzaro, N., & Mellon, L. (2005). *Fun meters for games*. In *Austin Game Developers Conference*.

Lewis-Beck, M., Bryman, A. E., & Liao, T. F. (2003). *The Sage encyclopedia of social science research methods*. Sage Publications.

Likert, R. (1932). A technique for the measurement of attitudes. *Archives of psychology*.

Livingston, I. J., Nacke, L. E., & Mandryk, R. L. (2011). Influencing experience: the effects of reading game reviews on player experience. In *Entertainment Computing–ICEC 2011* (pp. 89-100). Springer Berlin Heidelberg.

Malhotra, N., & Peterson, M. (2006). *Basic Marketing Research; a decision-making approach*. New Jersey, 7458.

McGonigal, J., 2011. *Technology is Broken*. The Penguin Press, New York.

Mekler, E. D., Bopp, J. A., Tuch, A. N., & Opwis, K. (2014, April). A systematic review of quantitative studies on the enjoyment of digital entertainment games. In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems* (pp. 927-936). ACM.

Metacritic (n.d.). How We Create the Metascore Magic. Retrieved August 29, 2015, <http://www.metacritic.com/about-metascores>.

Metacritic.com 2008, "About Metascores" <http://www.metacritic.com/about/scoring.shtml> EmSense, Inc., Corporate presentation

Nacke LE (2013) An introduction to physiological player metrics for evaluating games. In: Seif El-Nasr M, Drachen A, Canossa A (eds) *Game Analytics - Maximizing the Value of Player Data*. Springer London, 585–619. doi:10.1007/978-1-4471-4769-5\_26

Nacke LE, Bateman C, Mandryk RL (2014) BrainHex: A Neurobiological Gamer Typology Survey. *Entertain Comput* 5:55–62. doi:10.1016/j.entcom.2013.06.002

Nacke, L. E. (2015). *Games User Research and Physiological Game Evaluation*. In *Game User Experience Evaluation* (pp. 63-86). Springer International Publishing.

Nacke, L. E., & Lindley, C. A. (2010). Affective ludology, flow and immersion in a first-person shooter: Measurement of player experience. arXiv preprint arXiv:1004.0248.

Nordin, A. I., Denisova, A., & Cairns, P. (2014, October). Too Many Questionnaires: Measuring Player Experience Whilst Playing Digital Games. In *Seventh York Doctoral Symposium on Computer Science & Electronics* (p. 69).

Norman, K. L. Geq (game engagement/experience questionnaire): a review of two papers. *Interacting with Computers* 25(4) (2013), 278–283.

Qin, H., Rau, P.-L.P., Salvendy, G., 2009. Measuring player immersion in the computer game narrative. *International Journal of Human-Computer Interaction* 25 (2), 107–133.

Reichheld, F. F. (2003). The one number you need to grow. *Harvard business review*, 81(12), 46-55.

Reichheld, F. F. (2006). *The ultimate question: Driving good profits and true growth*. Harvard Business Press.

Ryan, R. M., Rigby, C. S., and Przybylski, A. K. Motivational pull of video games: A self-determination theory approach. *Motivation and Emotion* 30 (2006), 347–365.

Satmetrix Systems White Paper, <http://www.satmetrix.com/pdfs/netpromoterWPfinal.pdf>.

Satmetrix. (2004). *The power behind a single number: Growing your business with net promoter*.

Statista (2015) Breakdown of US Video game Sales 2009 by genre,

<http://www.statista.com/statistics/189592/breakdown-of-us-video-game-sales-2009-by-genre/>

Swain, C. in Isbister, K., & Schaffer, N. (2008). *Game usability: Advancing the player experience*. CRC Press.

Wyeth, P., Johnson, D. M., and Sweetser, P. Conceptualising, operationalising and measuring the player experience in videogames. In *Extended Proc. FnG '12*, IRIT Press (2012), 90–93.

## Appendix 1: Initial questionnaire

A few questions before we get started

What's your participant number?

What is your name?

What is your age?

What is your gender?

Please rate the following genres.

	Not at all interested	Somewhat interested	Very interested
First Person Shooter (FPS)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Adventure/Action-Adventure	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Platform	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Role-Playing Games (RPGs)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Sandbox (Open World)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Simulations	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Strategy/Tactics	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Sports	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Racing	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Party	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Appendix 2: Pre-game questionnaire

What's your participant number?

The game you'll be playing now is \_\_\_\_\_

Have you heard of this game before? YES / NO

Based on what you know about this game, what would you rate this game?

0	1	2	3	4	5	6	7	8	9	10
Poor					Average					Excellent



## Appendix 4: Immersive Experience Questionnaire (IEQ)

Please answer the following questions by circling the relevant number. In particular, remember that these questions are asking you about how you felt at the end of the game.

1. To what extent did the game hold your attention?  
Not at all    1    2    3    4    5    6    7    A lot
2. To what extent did you feel you were focused on the game?  
Not at all    1    2    3    4    5    6    7    A lot
3. How much effort did you put into playing the game?  
Very little    1    2    3    4    5    6    7    A lot
4. Did you feel that you were trying your best?  
Not at all    1    2    3    4    5    6    7    Very much so
5. To what extent did you lose track of time, e.g. did the game absorb your attention so that you were not bored?  
Not at all    1    2    3    4    5    6    7    A lot
6. To what extent did you feel consciously aware of being in the real world whilst playing?  
Not at all    1    2    3    4    5    6    7    Very much so
7. To what extent did you forget about your everyday concerns?  
Not at all    1    2    3    4    5    6    7    A lot
8. To what extent were you aware of yourself in your surroundings?  
Not at all    1    2    3    4    5    6    7    Very aware
9. To what extent did you notice events taking place around you?  
Not at all    1    2    3    4    5    6    7    A lot
10. Did you feel the urge at any point to stop playing and see what was happening around you?  
Not at all    1    2    3    4    5    6    7    Very much so
11. To what extent did you feel that you were interacting with the game environment?  
Not at all    1    2    3    4    5    6    7    Very much so
12. To what extent did you feel as though you were separated from your real-world environment?  
Not at all    1    2    3    4    5    6    7    Very much so
13. To what extent did you feel that the game was something fun you were experiencing, rather than a task you were just doing?  
Not at all    1    2    3    4    5    6    7    Very much so
14. To what extent was your sense of being in the game environment stronger than your sense of being in the real world?  
Not at all    1    2    3    4    5    6    7    Very much so

15. At any point did you find yourself become so involved that you were unaware you were even using controls, e.g. it was effortless?  
Not at all    1    2    3    4    5    6    7    Very much so
16. To what extent did you feel as though you were moving through the game according to your own will?  
Not at all    1    2    3    4    5    6    7    Very much so
17. To what extent did you find the game challenging?  
Not at all    1    2    3    4    5    6    7    Very difficult
18. Were there any times during the game in which you just wanted to give up?  
Not at all    1    2    3    4    5    6    7    A lot
19. To what extent did you feel motivated while playing?  
Not at all    1    2    3    4    5    6    7    A lot
20. To what extent did you find the game easy?  
Not at all    1    2    3    4    5    6    7    Very much so
21. To what extent did you feel like you were making progress towards the end of the game?  
Not at all    1    2    3    4    5    6    7    A lot
22. How well do you think you performed in the game?  
Very poor    1    2    3    4    5    6    7    Very well
23. To what extent did you feel emotionally attached to the game?  
Not at all    1    2    3    4    5    6    7    Very much so
24. To what extent were you interested in seeing how the game's events would progress?  
Not at all    1    2    3    4    5    6    7    A lot
25. How much did you want to "win" the game?  
Not at all    1    2    3    4    5    6    7    Very much so
26. Were you in suspense about whether or not you would do well in the game?  
Not at all    1    2    3    4    5    6    7    Very much so
27. At any point did you find yourself become so involved that you wanted to speak to the game directly?  
Not at all    1    2    3    4    5    6    7    Very much so
28. To what extent did you enjoy the graphics and the imagery?  
Not at all    1    2    3    4    5    6    7    A lot
29. How much would you say you enjoyed playing the game?  
Not at all    1    2    3    4    5    6    7    A lot
30. When it ended, were you disappointed that the game was over?  
Not at all    1    2    3    4    5    6    7    Very much so
31. Would you like to play the game again?  
Definitely no    1    2    3    4    5    6    7    Definitely yes

## Appendix 5: Game Experience Questionnaire (GEQ)

Please indicate how you felt while playing the game for each of the items, on the following scale:

0	1	2	3	4
not at all	slightly	moderately	fairly	extremely

- |  |  |
|--|--|
| 1 I felt content                       | 22 I enjoyed it                              |
| 2 I felt skilful                       | 23 I was fast at reaching the game's targets |
| 3 I was interested in the game's story | 24 I felt annoyed                            |
| 4 I could laugh about it               | 25 I was distracted                          |
| 5 I felt completely absorbed           | 26 I felt stimulated                         |
| 6 I felt happy                         | 27 I felt irritable                          |
| 7 I felt tense                         | 28 I lost track of time                      |
| 8 I felt that I was learning           | 29 I felt challenged                         |
| 9 I felt restless                      | 30 I found it impressive                     |
| 10 I thought about other things        | 31 I was deeply concentrated in the game     |
| 11 I found it tiresome                 | 32 I felt frustrated                         |
| 12 I felt strong                       | 33 It felt like a rich experience            |
| 13 I thought it was hard               | 34 I lost connection with the outside world  |
| 14 It was aesthetically pleasing       | 35 I was bored by the story                  |
| 15 I forgot everything around me       | 36 I had to put a lot of effort into it      |
| 16 I felt good                         | 37 I felt time pressure                      |
| 17 I was good at it                    | 38 It gave me a bad mood                     |
| 18 I felt bored                        | 39 I felt pressured                          |
| 19 I felt successful                   | 40 I was fully occupied with the game        |
| 20 I felt imaginative                  | 41 I thought it was fun                      |
| 21 I felt that I could explore things  | 42 I felt competent                          |



## Appendix 7: Post-game questionnaire

Order the five games from your most favourite to least favourite.

Order the five games from most immersive to least immersive.

Order the five games from best game experience to worst game experience.

Order the five games from the one you most likely to play again to the least likely to play again.

Order the five games from the one you most likely to recommend to your friends and colleagues to the least.

Order the five games from the most fun to the least fun.

You can drag and drop the games to your preferred order.

- Homefront
- Borderlands 2
- Metro Last Light
- Farcry 3
- Medal of Honor: War fighter

Appendix 8: Photos of Lab set-up



10-seat game testing area



Set-up in each pod

Appendix 9: Instruction sheet for each game

## Far Cry 3

### The mission you will be playing

The Medusa's Call, Playing The Spoiler

### Background

Jason came to the Rook Islands while on a skydiving trip with a group of friends and was captured by Pirates. Following his near-fatal escape from the insane Vaas Montenegro, Jason grew to become a legend among the residents of Rook Island and with the aid of the Rakyat, Jason becomes unstoppable and doesn't rest until his friends and family are rescued.

### Please pause the game and fill in the questionnaires after...

Escape the area or kill all remaining pirates nearby



# Medal of Honor: Warfighter

## The mission you will be playing

Changing Tides, Rip Current

## Background

Medal of Honor Warfighter tells the story of U.S. Tier 1 Operator, "Preacher" as he returns home from overseas only to find his family torn apart from years of deployment. Trying to pick up the pieces to salvage what remains of his marriage, Preacher is reminded of what he's fighting for - family. But when an extremely deadly explosive (PETN) penetrates civilian borders and his two worlds collide, Preacher and his fellow teammates are sent in to solve the problem. They take the fight to the enemy and do whatever it takes to protect their loved ones from harm.

## Please pause the game and fill in the questionnaires when

....

you have reached the boat.



# Borderlands 2

## The mission you will be playing

Best Minion Ever

## Background

A new era of shoot and loot is about to begin. You will be playing as a vault hunter facing off against a massive new world of creatures, psychos and the evil mastermind, Handsome Jack. Unfortunately, one man stands between you and your next destination. His name: Captain Flynt.

## Please pause the game and fill in the questionnaires when

....

The boss, Captain Flynt is dead and you are continuing on to Claptrap's ship



# Metro Last Light

## The mission you will be playing

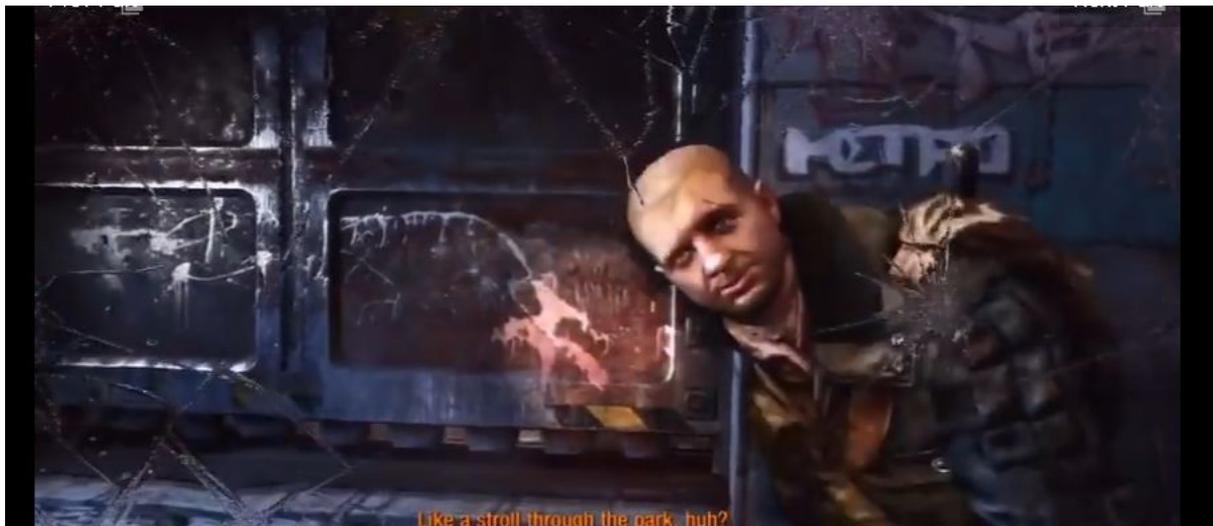
Facilities, Torchlight, Echo

## Background

In the year of 2033, Artyom journeys across linear sections of a Post-Apocalyptic Metro System and an irradiated surface of post-apocalyptic Moscow, hoping to save what's left of humanity, long after World War III. In this inhospitable landscape, you must guide Artyom into following friends, battling dangerous mutants, and engaging in firefights with enemy factions, or complex stealth missions in which remaining unseen is vital for survival.

## Please pause the game and fill in the questionnaires after...

you kill the monsters on the escalators, then join Pavel back in the Metro.



## Homefront

## The mission you will be playing

Why We Fight, Freedom

## Background

Welcome to 2027. It's been two years since the KPA attacked the United States. You are a former Marine pilot just trying to keep a low profile. A week ago, you received a draft notice; the Koreans want to recruit you. But they're not the only ones.

**Please pause the game and fill in the questionnaires after...**

you clear out the remaining Koreans.



## Appendix 10: Information sheet for participants

# INFORMATION SHEET FOR PARTICIPANTS

Thank you for taking part in this experiment.

Today you will be playing 5 games, and after each game you will fill in a questionnaire. Please play these 5 games as if you were at home and just reached this level.

Before the gameplay experience there is a demographics questionnaire for you to fill in, and after playing all 5 games you will need to fill in a final questionnaire.

Below is a more detailed plan for today.

1. Game 1 (60 min)
  - i. Fills out sets of questionnaire on tablet screen (20 min)
  - ii. Break (10 min)
2. Game 2 (60 min)
  - i. Fills out sets of questionnaire on tablet screen (20 min)
3. Lunch (60 min)
4. Game 3 (60 min)
  - i. Fills out sets of questionnaire on tablet screen (20 min)
  - ii. Break (10 min)
5. Game 4 (60 min)
  - i. Fills out sets of questionnaire on tablet screen (20 min)
  - ii. Break (10 min)
6. Game 5 (60 min)
  - i. Fills out sets of questionnaire on tablet screen (20 min)
7. Fills in final questionnaire (10 min)
8. Debriefs; Receives compensation

## Appendix 11: Informed consent form

# INFORMED CONSENT FORM

Title of Project: First Person Shooter Game Experience

This study has been approved by the UCL Research Ethics Committee [Project ID Number]:  
UCLIC/MS/1112/002/Cox/CitizenCyberlab

We are investigating the game experience in First Person Shooter (FPS) games.

In this experiment, you will be asked to play a selected chapter of each of the five games, followed by a questionnaire after each. The experiment will last for about 8 hours. You will be playing each of the games for an hour, and with short breaks in between. Complimentary lunch and refreshments will be provided.

The records of this study will be kept confidential in accordance with the Data Protection Act 1998, only researchers will have access to the records. We will not include any information that makes it possible to identify you in any report. Please be aware that you are free to withdraw from the study without penalty at any point of the experiment.

Please ask any questions you have now, or at any point during the experiment.

Contact details:

Christie Lau      christie.lau.09@ucl.ac.uk

### Participant's Statement

I

- have read the notes written above and the Information Sheet, and understand what the study involves.
- understand that if I decide at any time that I no longer wish to take part in this project, I can notify the researches involved and withdraw immediately.
- consent to the processing of my personal information for the purposes of this research study.
- understand that such information will be treated as strictly confidential and handled in accordance with the provisions of the Data Protection Act 1998.
- agree that the research project named above has been explained to my satisfaction and I agree to take part in this study.
- understand that the information I have submitted will be published as a report and I may request a copy. Confidentiality and anonymity will be maintained and it will not be possible to identify me from any publications.
- agree that my non-personal research data may be used by others for future research. I am assured that the confidentiality of my personal data will be upheld through the removal of identities.
- agree to not discuss this study with others today and for about a week after today, since the study is ongoing.

Signed:

Name in block letters:

Date: July, 2015

### Investigator's Statement

I, Christie Lau, confirm that I have carefully explained the purpose of the study to the participant and outlined any reasonably foreseeable risks or benefits (where applicable).

Signed:

Date: July, 2015

APPENDIX 12 Latin Square for Play Sequence

Note: The number after the game name is the assigned pod number

Time slot	Participant 1, 11, 21, 31	Participant 2, 12, 22, 32	Participant 3, 13, 23, 33	Participant 4, 14, 24, 34	Participant 5, 15, 25, 35
1	MoH 7	Homefront 9	Far cry 5	Metro 1	Borderlands 3
2	Borderlands 3	Far cry 5	MoH 7	Homefront 9	Metro 1
3	Far cry 5	MoH 7	Metro 1	Borderlands 3	Homefront 9
4	Metro 1	Borderlands 3	Homefront 9	Far cry 5	MoH 7
5	Homefront 9	Metro 1	Borderlands 3	MoH 7	Far cry 5

Time slot	Participant 6, 16, 26, 36	Participant 7, 17, 27, 37	Participant 18, 28, 38	Participant 9, 19, 29, 39	Participant 10, 20, 30, 40
1	MoH 8	Homefront 10	Far cry 6	Metro 2	Borderlands 4
2	Borderlands 4	Far cry 6	MoH 8	Homefront 10	Metro 2
3	Far cry 6	MoH 8	Metro 2	Borderlands 4	Homefront 10
4	Metro 2	Borderlands 4	Homefront 10	Far cry 6	MoH 8
5	Homefront 10	Metro 2	Borderlands 4	MoH 8	Far cry 6

APPENDIX 13 Full table of Post hoc test using the Bonferroni correction

Measure	(I) Games	(J) Games	Mean Difference (I-J)	Std. Error	Sig.b
GEQ Competence	Borderlands 2	Farcry 3	-0.026	0.173	1
		Metro Last Light	-0.497	0.194	0.142
		Homefront	-0.221	0.171	1
		Medal of Honor: War Fighter	-0.179	0.177	1
	Farcry 3	Borderlands 2	0.026	0.173	1
		Metro Last Light	-0.472	0.182	0.133
		Homefront	-0.195	0.162	1
		Medal of Honor: War Fighter	-0.154	0.17	1
	Metro Last Light	Borderlands 2	0.497	0.194	0.142
		Farcry 3	0.472	0.182	0.133
		Homefront	0.277	0.129	0.378
		Medal of Honor: War Fighter	0.318	0.176	0.78
	Homefront	Borderlands 2	0.221	0.171	1
		Farcry 3	0.195	0.162	1
		Metro Last Light	-0.277	0.129	0.378
		Medal of Honor: War Fighter	0.041	0.176	1
	Medal of Honor: War Fighter	Borderlands 2	0.179	0.177	1
		Farcry 3	0.154	0.17	1
		Metro Last Light	-0.318	0.176	0.78
		Homefront	-0.041	0.176	1
GEQ_Immersion	Borderlands 2	Farcry 3	-0.081	0.221	1
		Metro Last Light	0.021	0.252	1
		Homefront	0.226	0.239	1
		Medal of Honor:	-.782*	0.178	0.001

		War Fighter			
	Farcry 3	Borderlands 2	0.081	0.221	1
		Metro Last Light	0.103	0.175	1
		Homefront	0.308	0.194	1
		Medal of Honor: War Fighter	-.701*	0.179	0.004
	Metro Last Light	Borderlands 2	-0.021	0.252	1
		Farcry 3	-0.103	0.175	1
		Homefront	0.205	0.149	1
		Medal of Honor: War Fighter	-.803*	0.187	0.001
	Homefront	Borderlands 2	-0.226	0.239	1
		Farcry 3	-0.308	0.194	1
		Metro Last Light	-0.205	0.149	1
		Medal of Honor: War Fighter	-1.009*	0.182	0
	Medal of Honor: War Fighter	Borderlands 2	.782*	0.178	0.001
		Farcry 3	.701*	0.179	0.004
		Metro Last Light	.803*	0.187	0.001
		Homefront	1.009*	0.182	0
GEQ Flow	Borderlands 2	Farcry 3	0.067	0.232	1
		Metro Last Light	-0.103	0.225	1
		Homefront	0.221	0.208	1
		Medal of Honor: War Fighter	-0.087	0.201	1
	Farcry 3	Borderlands 2	-0.067	0.232	1
		Metro Last Light	-0.169	0.18	1
		Homefront	0.154	0.187	1
		Medal of Honor: War Fighter	-0.154	0.217	1
	Metro Last Light	Borderlands 2	0.103	0.225	1
		Farcry 3	0.169	0.18	1

		Homefront	0.323	0.156	0.452
		Medal of Honor: War Fighter	0.015	0.209	1
	Homefront	Borderlands 2	-0.221	0.208	1
		Farcry 3	-0.154	0.187	1
		Metro Last Light	-0.323	0.156	0.452
		Medal of Honor: War Fighter	-0.308	0.18	0.962
	Medal of Honor: War Fighter	Borderlands 2	0.087	0.201	1
		Farcry 3	0.154	0.217	1
		Metro Last Light	-0.015	0.209	1
		Homefront	0.308	0.18	0.962
GEQ Tension	Borderlands 2	Farcry 3	-0.041	0.199	1
		Metro Last Light	0.005	0.172	1
		Homefront	0.174	0.142	1
		Medal of Honor: War Fighter	0.046	0.2	1
	Farcry 3	Borderlands 2	0.041	0.199	1
		Metro Last Light	0.046	0.18	1
		Homefront	0.215	0.182	1
		Medal of Honor: War Fighter	0.087	0.168	1
	Metro Last Light	Borderlands 2	-0.005	0.172	1
		Farcry 3	-0.046	0.18	1
		Homefront	0.169	0.173	1
		Medal of Honor: War Fighter	0.041	0.158	1
	Homefront	Borderlands 2	-0.174	0.142	1
		Farcry 3	-0.215	0.182	1
		Metro Last Light	-0.169	0.173	1
		Medal of Honor: War Fighter	-0.128	0.21	1

	Medal of Honor: War Fighter	Borderlands 2	-0.046	0.2	1
		Farcry 3	-0.087	0.168	1
		Metro Last Light	-0.041	0.158	1
		Homefront	0.128	0.21	1
GEQ Challenge	Borderlands 2	Farcry 3	0.118	0.171	1
		Metro Last Light	0	0.152	1
		Homefront	0.415	0.162	0.143
		Medal of Honor: War Fighter	-0.215	0.116	0.723
	Farcry 3	Borderlands 2	-0.118	0.171	1
		Metro Last Light	-0.118	0.177	1
		Homefront	0.297	0.155	0.622
		Medal of Honor: War Fighter	-0.333	0.18	0.714
	Metro Last Light	Borderlands 2	0	0.152	1
		Farcry 3	0.118	0.177	1
		Homefront	0.415	0.146	0.073
		Medal of Honor: War Fighter	-0.215	0.161	1
	Homefront	Borderlands 2	-0.415	0.162	0.143
		Farcry 3	-0.297	0.155	0.622
		Metro Last Light	-0.415	0.146	0.073
		Medal of Honor: War Fighter	-.631*	0.163	0.004
	Medal of Honor: War Fighter	Borderlands 2	0.215	0.116	0.723
		Farcry 3	0.333	0.18	0.714
		Metro Last Light	0.215	0.161	1
		Homefront	.631*	0.163	0.004
GEQ Negative Affect	Borderlands 2	Farcry 3	0.2	0.195	1
		Metro Last Light	0.128	0.196	1
		Homefront	0.313	0.192	1

		Medal of Honor: War Fighter	0.338	0.177	0.631
	Farcry 3	Borderlands 2	-0.2	0.195	1
		Metro Last Light	-0.072	0.183	1
		Homefront	0.113	0.18	1
		Medal of Honor: War Fighter	0.138	0.179	1
	Metro Last Light	Borderlands 2	-0.128	0.196	1
		Farcry 3	0.072	0.183	1
		Homefront	0.185	0.151	1
		Medal of Honor: War Fighter	0.21	0.184	1
	Homefront	Borderlands 2	-0.313	0.192	1
		Farcry 3	-0.113	0.18	1
		Metro Last Light	-0.185	0.151	1
		Medal of Honor: War Fighter	0.026	0.175	1
	Medal of Honor: War Fighter	Borderlands 2	-0.338	0.177	0.631
		Farcry 3	-0.138	0.179	1
		Metro Last Light	-0.21	0.184	1
		Homefront	-0.026	0.175	1
GEQ Positive Affect	Borderlands 2	Farcry 3	0.041	0.23	1
		Metro Last Light	-0.103	0.236	1
		Homefront	-0.149	0.21	1
		Medal of Honor: War Fighter	-0.277	0.207	1
	Farcry 3	Borderlands 2	-0.041	0.23	1
		Metro Last Light	-0.144	0.195	1
		Homefront	-0.19	0.181	1
		Medal of Honor: War Fighter	-0.318	0.205	1
	Metro Last Light	Borderlands 2	0.103	0.236	1

		Farcry 3	0.144	0.195	1
		Homefront	-0.046	0.151	1
		Medal of Honor: War Fighter	-0.174	0.199	1
	Homefront	Borderlands 2	0.149	0.21	1
		Farcry 3	0.19	0.181	1
		Metro Last Light	0.046	0.151	1
		Medal of Honor: War Fighter	-0.128	0.164	1
	Medal of Honor: War Fighter	Borderlands 2	0.277	0.207	1
		Farcry 3	0.318	0.205	1
		Metro Last Light	0.174	0.199	1
		Homefront	0.128	0.164	1
How would you RATE the game?	Borderlands 2	Farcry 3	-0.205	0.544	1
		Metro Last Light	-0.41	0.535	1
		Homefront	-0.179	0.462	1
		Medal of Honor: War Fighter	-1.154	0.398	0.062
	Farcry 3	Borderlands 2	0.205	0.544	1
		Metro Last Light	-0.205	0.418	1
		Homefront	0.026	0.455	1
		Medal of Honor: War Fighter	-0.949	0.45	0.416
	Metro Last Light	Borderlands 2	0.41	0.535	1
		Farcry 3	0.205	0.418	1
		Homefront	0.231	0.327	1
		Medal of Honor: War Fighter	-0.744	0.425	0.881
	Homefront	Borderlands 2	0.179	0.462	1
		Farcry 3	-0.026	0.455	1
		Metro Last Light	-0.231	0.327	1
		Medal of Honor: War Fighter	-0.974	0.361	0.103

		War Fighter			
	Medal of Honor: War Fighter	Borderlands 2	1.154	0.398	0.062
		Farcry 3	0.949	0.45	0.416
		Metro Last Light	0.744	0.425	0.881
		Homefront	0.974	0.361	0.103
How would you rate the GRAPHICS in the game?	Borderlands 2	Farcry 3	-0.795	0.545	1
		Metro Last Light	-0.949	0.518	0.75
		Homefront	0.641	0.484	1
		Medal of Honor: War Fighter	-1.718*	0.422	0.002
	Farcry 3	Borderlands 2	0.795	0.545	1
		Metro Last Light	-0.154	0.443	1
		Homefront	1.436*	0.457	0.032
		Medal of Honor: War Fighter	-0.923	0.458	0.509
	Metro Last Light	Borderlands 2	0.949	0.518	0.75
		Farcry 3	0.154	0.443	1
		Homefront	1.590*	0.481	0.021
		Medal of Honor: War Fighter	-0.769	0.393	0.577
	Homefront	Borderlands 2	-0.641	0.484	1
		Farcry 3	-1.436*	0.457	0.032
		Metro Last Light	-1.590*	0.481	0.021
		Medal of Honor: War Fighter	-2.359*	0.465	0
	Medal of Honor: War Fighter	Borderlands 2	1.718*	0.422	0.002
		Farcry 3	0.923	0.458	0.509
		Metro Last Light	0.769	0.393	0.577
		Homefront	2.359*	0.465	0
How would you rate the SOUND	Borderlands 2	Farcry 3	-0.179	0.501	1
		Metro Last Light	-0.154	0.473	1

EFFECTS in the game?		Homefront	0.103	0.434	1
		Medal of Honor: War Fighter	-0.641	0.338	0.653
	Farcry 3	Borderlands 2	0.179	0.501	1
		Metro Last Light	0.026	0.527	1
		Homefront	0.282	0.45	1
		Medal of Honor: War Fighter	-0.462	0.451	1
	Metro Last Light	Borderlands 2	0.154	0.473	1
		Farcry 3	-0.026	0.527	1
		Homefront	0.256	0.426	1
		Medal of Honor: War Fighter	-0.487	0.414	1
	Homefront	Borderlands 2	-0.103	0.434	1
		Farcry 3	-0.282	0.45	1
		Metro Last Light	-0.256	0.426	1
		Medal of Honor: War Fighter	-0.744	0.348	0.391
	Medal of Honor: War Fighter	Borderlands 2	0.641	0.338	0.653
		Farcry 3	0.462	0.451	1
Metro Last Light		0.487	0.414	1	
Homefront		0.744	0.348	0.391	
How would you rate the MUSIC of the game?	Borderlands 2	Farcry 3	0.051	0.5	1
		Metro Last Light	0.282	0.457	1
		Homefront	0.205	0.482	1
		Medal of Honor: War Fighter	-0.744	0.414	0.801
	Farcry 3	Borderlands 2	-0.051	0.5	1
		Metro Last Light	0.231	0.353	1
		Homefront	0.154	0.37	1
		Medal of Honor: War Fighter	-0.795	0.399	0.538

	Metro Last Light	Borderlands 2	-0.282	0.457	1
		Farcry 3	-0.231	0.353	1
		Homefront	-0.077	0.375	1
		Medal of Honor: War Fighter	-1.026	0.393	0.129
	Homefront	Borderlands 2	-0.205	0.482	1
		Farcry 3	-0.154	0.37	1
		Metro Last Light	0.077	0.375	1
		Medal of Honor: War Fighter	-0.949	0.352	0.105
	Medal of Honor: War Fighter	Borderlands 2	0.744	0.414	0.801
		Farcry 3	0.795	0.399	0.538
		Metro Last Light	1.026	0.393	0.129
		Homefront	0.949	0.352	0.105
	Overall, I ENJOYED playing the game	Borderlands 2	Farcry 3	-0.436	0.591
Metro Last Light			-0.59	0.581	1
Homefront			-0.385	0.514	1
Medal of Honor: War Fighter			-1.308*	0.404	0.025
Farcry 3		Borderlands 2	0.436	0.591	1
		Metro Last Light	-0.154	0.451	1
		Homefront	0.051	0.482	1
		Medal of Honor: War Fighter	-0.872	0.489	0.826
Metro Last Light		Borderlands 2	0.59	0.581	1
		Farcry 3	0.154	0.451	1
		Homefront	0.205	0.329	1
		Medal of Honor: War Fighter	-0.718	0.431	1
Homefront		Borderlands 2	0.385	0.514	1
		Farcry 3	-0.051	0.482	1
		Metro Last Light	-0.205	0.329	1

		Medal of Honor: War Fighter	-0.923	0.349	0.119
	Medal of Honor: War Fighter	Borderlands 2	1.308*	0.404	0.025
		Farcry 3	0.872	0.489	0.826
		Metro Last Light	0.718	0.431	1
		Homefront	0.923	0.349	0.119
NET promoter score	Borderlands 2	Farcry 3	-0.538	0.664	1
		Metro Last Light	-0.769	0.696	1
		Homefront	-0.308	0.662	1
		Medal of Honor: War Fighter	-1.385	0.531	0.129
	Farcry 3	Borderlands 2	0.538	0.664	1
		Metro Last Light	-0.231	0.519	1
		Homefront	0.231	0.563	1
		Medal of Honor: War Fighter	-0.846	0.571	1
	Metro Last Light	Borderlands 2	0.769	0.696	1
		Farcry 3	0.231	0.519	1
		Homefront	0.462	0.473	1
		Medal of Honor: War Fighter	-0.615	0.592	1
	Homefront	Borderlands 2	0.308	0.662	1
		Farcry 3	-0.231	0.563	1
		Metro Last Light	-0.462	0.473	1
		Medal of Honor: War Fighter	-1.077	0.526	0.478
	Medal of Honor: War Fighter	Borderlands 2	1.385	0.531	0.129
		Farcry 3	0.846	0.571	1
		Metro Last Light	0.615	0.592	1
		Homefront	1.077	0.526	0.478
IEQ_Total	Borderlands 2	Farcry 3	-4.667	7.591	1
		Metro Last Light	-7.359	8.318	1

		Homefront	1.821	7.153	1
		Medal of Honor: War Fighter	-11.923	6.905	0.923
	Farcry 3	Borderlands 2	4.667	7.591	1
		Metro Last Light	-2.692	6.408	1
		Homefront	6.487	6.226	1
		Medal of Honor: War Fighter	-7.256	6.69	1
	Metro Last Light	Borderlands 2	7.359	8.318	1
		Farcry 3	2.692	6.408	1
		Homefront	9.179	5.296	0.911
		Medal of Honor: War Fighter	-4.564	6.374	1
	Homefront	Borderlands 2	-1.821	7.153	1
		Farcry 3	-6.487	6.226	1
		Metro Last Light	-9.179	5.296	0.911
		Medal of Honor: War Fighter	-13.744	5.66	0.2
		Borderlands 2	11.923	6.905	0.923
	Medal of Honor: War Fighter	Farcry 3	7.256	6.69	1
		Metro Last Light	4.564	6.374	1
		Homefront	13.744	5.66	0.2
IEQ_CI	Borderlands 2	Farcry 3	-0.667	2.588	1
		Metro Last Light	-1.923	2.874	1
		Homefront	-0.179	2.501	1
		Medal of Honor: War Fighter	-2.179	2.196	1
	Farcry 3	Borderlands 2	0.667	2.588	1
		Metro Last Light	-1.256	2.338	1
		Homefront	0.487	1.88	1
		Medal of Honor: War Fighter	-1.513	1.92	1

	Metro Last Light	Borderlands 2	1.923	2.874	1
		Farcry 3	1.256	2.338	1
		Homefront	1.744	1.792	1
		Medal of Honor: War Fighter	-0.256	2.22	1
	Homefront	Borderlands 2	0.179	2.501	1
		Farcry 3	-0.487	1.88	1
		Metro Last Light	-1.744	1.792	1
		Medal of Honor: War Fighter	-2	1.926	1
	Medal of Honor: War Fighter	Borderlands 2	2.179	2.196	1
		Farcry 3	1.513	1.92	1
		Metro Last Light	0.256	2.22	1
		Homefront	2	1.926	1
	IEQ_RWD	Borderlands 2	Farcry 3	-0.103	1.391
Metro Last Light			-1.256	1.509	1
Homefront			1.333	1.128	1
Medal of Honor: War Fighter			-1.436	1.63	1
Farcry 3		Borderlands 2	0.103	1.391	1
		Metro Last Light	-1.154	1.163	1
		Homefront	1.436	1.281	1
		Medal of Honor: War Fighter	-1.333	1.528	1
Metro Last Light		Borderlands 2	1.256	1.509	1
		Farcry 3	1.154	1.163	1
		Homefront	2.59	1.177	0.339
		Medal of Honor: War Fighter	-0.179	1.339	1
Homefront		Borderlands 2	-1.333	1.128	1
		Farcry 3	-1.436	1.281	1
		Metro Last Light	-2.59	1.177	0.339

		Medal of Honor: War Fighter	-2.769	1.067	0.133
	Medal of Honor: War Fighter	Borderlands 2	1.436	1.63	1
		Farcry 3	1.333	1.528	1
		Metro Last Light	0.179	1.339	1
		Homefront	2.769	1.067	0.133
IEQ_EI	Borderlands 2	Farcry 3	-4.59	3.46	1
		Metro Last Light	-4.487	3.734	1
		Homefront	-2.795	3.397	1
		Medal of Honor: War Fighter	-6.487	2.935	0.332
	Farcry 3	Borderlands 2	4.59	3.46	1
		Metro Last Light	0.103	3.082	1
		Homefront	1.795	2.782	1
		Medal of Honor: War Fighter	-1.897	3.047	1
	Metro Last Light	Borderlands 2	4.487	3.734	1
		Farcry 3	-0.103	3.082	1
		Homefront	1.692	2.553	1
		Medal of Honor: War Fighter	-2	3.048	1
	Homefront	Borderlands 2	2.795	3.397	1
		Farcry 3	-1.795	2.782	1
		Metro Last Light	-1.692	2.553	1
		Medal of Honor: War Fighter	-3.692	2.779	1
	Medal of Honor: War Fighter	Borderlands 2	6.487	2.935	0.332
		Farcry 3	1.897	3.047	1
		Metro Last Light	2	3.048	1
		Homefront	3.692	2.779	1
IEQ_CHA	Borderlands 2	Farcry 3	-0.436	0.624	1
		Metro Last Light	0.256	0.651	1

		Homefront	-0.026	0.648	1
		Medal of Honor: War Fighter	-0.436	0.594	1
	Farcry 3	Borderlands 2	0.436	0.624	1
		Metro Last Light	0.692	0.616	1
		Homefront	0.41	0.638	1
		Medal of Honor: War Fighter	0	0.58	1
	Metro Last Light	Borderlands 2	-0.256	0.651	1
		Farcry 3	-0.692	0.616	1
		Homefront	-0.282	0.596	1
		Medal of Honor: War Fighter	-0.692	0.614	1
	Homefront	Borderlands 2	0.026	0.648	1
		Farcry 3	-0.41	0.638	1
		Metro Last Light	0.282	0.596	1
		Medal of Honor: War Fighter	-0.41	0.575	1
		Borderlands 2	0.436	0.594	1
	Medal of Honor: War Fighter	Farcry 3	0	0.58	1
		Metro Last Light	0.692	0.614	1
		Homefront	0.41	0.575	1
IEQ_CON	Borderlands 2	Farcry 3	-1.051	1.42	1
		Metro Last Light	-0.462	1.663	1
		Homefront	3.538	1.673	0.41
		Medal of Honor: War Fighter	-3.667	1.48	0.178
	Farcry 3	Borderlands 2	1.051	1.42	1
		Metro Last Light	0.59	1.341	1
		Homefront	4.59	1.754	0.127
		Medal of Honor: War Fighter	-2.615	1.629	1

	Metro Last Light	Borderlands 2	0.462	1.663	1
		Farcry 3	-0.59	1.341	1
		Homefront	4.000*	1.29	0.036
		Medal of Honor: War Fighter	-3.205	1.268	0.158
	Homefront	Borderlands 2	-3.538	1.673	0.41
		Farcry 3	-4.59	1.754	0.127
		Metro Last Light	-4.000*	1.29	0.036
		Medal of Honor: War Fighter	-7.205*	1.47	0
	Medal of Honor: War Fighter	Borderlands 2	3.667	1.48	0.178
		Farcry 3	2.615	1.629	1
		Metro Last Light	3.205	1.268	0.158
		Homefront	7.205*	1.47	0
PENS_AUT	Borderlands 2	Farcry 3	0.641	0.365	0.869
		Metro Last Light	1.085*	0.343	0.03
		Homefront	1.060*	0.305	0.013
		Medal of Honor: War Fighter	-1.521*	0.267	0
	Farcry 3	Borderlands 2	-0.641	0.365	0.869
		Metro Last Light	0.444	0.288	1
		Homefront	0.419	0.343	1
		Medal of Honor: War Fighter	-2.162*	0.28	0
	Metro Last Light	Borderlands 2	-1.085*	0.343	0.03
		Farcry 3	-0.444	0.288	1
		Homefront	-0.026	0.222	1
		Medal of Honor: War Fighter	-2.607*	0.267	0
	Homefront	Borderlands 2	-1.060*	0.305	0.013
		Farcry 3	-0.419	0.343	1
		Metro Last Light	0.026	0.222	1

		Medal of Honor: War Fighter	-2.581*	0.298	0
	Medal of Honor: War Fighter	Borderlands 2	1.521*	0.267	0
		Farcry 3	2.162*	0.28	0
		Metro Last Light	2.607*	0.267	0
		Homefront	2.581*	0.298	0
PENS_COM	Borderlands 2	Farcry 3	-0.043	0.316	1
		Metro Last Light	-0.427	0.293	1
		Homefront	-0.538	0.284	0.658
		Medal of Honor: War Fighter	-0.068	0.276	1
	Farcry 3	Borderlands 2	0.043	0.316	1
		Metro Last Light	-0.385	0.301	1
		Homefront	-0.496	0.286	0.91
		Medal of Honor: War Fighter	-0.026	0.253	1
	Metro Last Light	Borderlands 2	0.427	0.293	1
		Farcry 3	0.385	0.301	1
		Homefront	-0.111	0.199	1
		Medal of Honor: War Fighter	0.359	0.244	1
	Homefront	Borderlands 2	0.538	0.284	0.658
		Farcry 3	0.496	0.286	0.91
		Metro Last Light	0.111	0.199	1
		Medal of Honor: War Fighter	0.47	0.289	1
	Medal of Honor: War Fighter	Borderlands 2	0.068	0.276	1
		Farcry 3	0.026	0.253	1
		Metro Last Light	-0.359	0.244	1
		Homefront	-0.47	0.289	1
PENS_PRE	Borderlands 2	Farcry 3	-0.444	0.236	0.669
		Metro Last Light	-0.288	0.292	1

		Homefront	-0.35	0.246	1	
		Medal of Honor: War Fighter	-.752*	0.228	0.021	
Farcry 3		Borderlands 2	0.444	0.236	0.669	
		Metro Last Light	0.157	0.246	1	
		Homefront	0.094	0.23	1	
		Medal of Honor: War Fighter	-0.308	0.228	1	
Metro Last Light		Borderlands 2	0.288	0.292	1	
		Farcry 3	-0.157	0.246	1	
		Homefront	-0.063	0.171	1	
		Medal of Honor: War Fighter	-0.464	0.207	0.305	
Homefront		Borderlands 2	0.35	0.246	1	
		Farcry 3	-0.094	0.23	1	
		Metro Last Light	0.063	0.171	1	
		Medal of Honor: War Fighter	-0.402	0.182	0.332	
Medal of Honor: War Fighter		Borderlands 2	.752*	0.228	0.021	
		Farcry 3	0.308	0.228	1	
		Metro Last Light	0.464	0.207	0.305	
		Homefront	0.402	0.182	0.332	
PENS_IC	Borderlands 2	Farcry 3	0.017	0.257	1	
		Metro Last Light	-.769*	0.242	0.03	
		Homefront	-0.709	0.24	0.054	
		Medal of Honor: War Fighter	-0.137	0.257	1	
	Farcry 3		Borderlands 2	-0.017	0.257	1
			Metro Last Light	-.786*	0.259	0.043
			Homefront	-0.726	0.286	0.154
			Medal of Honor: War Fighter	-0.154	0.196	1

	Metro Last Light	Borderlands 2	.769*	0.242	0.03
		Farcry 3	.786*	0.259	0.043
		Homefront	0.06	0.195	1
		Medal of Honor: War Fighter	0.632	0.221	0.069
	Homefront	Borderlands 2	0.709	0.24	0.054
		Farcry 3	0.726	0.286	0.154
		Metro Last Light	-0.06	0.195	1
		Medal of Honor: War Fighter	0.573	0.271	0.411
	Medal of Honor: War Fighter	Borderlands 2	0.137	0.257	1
		Farcry 3	0.154	0.196	1
		Metro Last Light	-0.632	0.221	0.069
		Homefront	-0.573	0.271	0.411