RUN FOR FUN
MEASURING EXERGAME ENJOYMENT

An Interactive Qualifying Project
submitted to the faculty of the
WORCESTER POLYTECHNIC INSTITUTE
in partial fulfillment of the requirements for the
Degree of Bachelor of Science

AUTHORS

Alexander Fitzgerald

Sam Huang

Kyle Sposato

Dongjie Wang

ADVISORS

Professor Emmanuel Agu

Professor Mark Claypool

Table of Contents

1 Introduction **3**

2 Background and Exergame Literature Review **5**

2.1 Questionnaires

2.1.1 Game Engagement Questionnaire **6**

2.1.1.1 Game Engagement Questionnaire Critique **8**

2.1.2 Immersive Experience Questionnaire **12**

2.1.2.1 Immersive Experience Questionnaire Critique **13**

2.1.3 Physical Activity Enjoyment Scale Questionnaire **15**

2.1.4 Other Questionnaires **16**

2.2 Understanding Exergame Enjoyment **19**

2.2.1 Defining Exergame Enjoyment

2.2.2 Components to Test **20**

2.3 EEQ Considerations **21**

2.3.1 Conducting and Budgeting

2.3.2 How to Ask Questions and Mistakes to Avoid **22**

2.4 Exergame Research **24**

3 Methodology **26**

3.1 Focus Groups **27**

3.2 Participant Distribution

3.3 Exergame Procedures **30**

3.4 Questionnaires **32**

3.5 Scoring **33**

4 Exergame Choices **34**

4.1 Game choice criteria **35**

4.2 Top Fifteen Exergames and Elimination Rationales **37**

4.3 Exergames Selected for Evaluation **39**

4.4 Exergame Procedures **42**

5 Questionnaire Development **43**
5.1 Questions from Existing Questionnaires **44**
5.2 Questions Inspired by the PACES Questionnaire **47**
5.3 Original Questions **49**
5.4 The Exergame Enjoyment Questionnaire **52**
5.4.1 Categories **53**
5.4.2 Format
5.5 Additional Questionnaires **56**
6 Focus Groups **58**
6.1 Procedure
6.2 Methodology Revisions **60**
6.3 Questionnaire Revisions **63**
7 Questionnaire Validation **66**
7.1 Coding Procedure **67**
7.2 Validation Results **72**
7.3 Validation Discussion **81**
8 Evaluating Exergame Enjoyment using the EEQ **82**
8.1 Pokémon GO Enjoyment **83**
8.2 Just Dance Now! Enjoyment **86**
8.3 Enjoyment Discussion **89**
9 Conclusion and Future Work **92**
9.1 Questionnaire Validation Results **93**
9.2 Questionnaire Discussion **95**
9.3 Methodology Discussion **98**
9.4 Final Observations **100**
9.5 Recommendations for Future Work **102**
References **103**
Appendices **105**

Chapter 1 Introduction

Physical inactivity has become an increasingly relevant concern regarding the health of young adults. Many people are lacking the exercise needed to be healthy. Exercise is defined as, "activity requiring physical effort, carried out to sustain or improve health and fitness" (oxforddictionaries.com). A sedentary lifestyle poses an "Increased risk of a variety of health issues including coronary heart disease, metabolic syndrome, type 2 diabetes, breast and colon cancer, depression and falling" (McVeigh, 2016). Unfortunately, young adults who do not exercise choose instead to find satisfaction from electronic media. The use of mobile devices by this demographic, and technology in general, has also been increasing and is generally not conducive to exercise. Thus, methods of exercise that are delivered on mobile devices would be effective for a young demographic.

The emergence of mobile "exergames" provides a new method of exercise for people with sedentary lifestyles. Exergames are defined as, "A video game that also provides the player with physical exercise" (yourdictionary.com). Exergames can benefit players through a range of activity levels. For example, *Just Dance Now* (Ubisoft, 2014) is a game that provides physical activity in the form of dancing. *Pokémon GO* (Niantic Labs, 2016) requires players to walk to certain locations to catch Pokémon and gather items.

By "gamifying" exercise, young adults are encouraged to engage in physical activity while playing games. Exergames can provide regular exercise, as they are, "Enjoyable tools that increase energy expenditure during play, motivate players to become more physically active, promote social interaction, and enhance cognitive performance" (Staiano, 2011). Additionally, Exergames also provide biometric feedback such as: calories burned, heart rate, and steps taken as a means of measuring physical activity.

By utilizing the physical activity provided by exergames, it is possible to address the problem of physical inactivity in young adults. However, it has become necessary to develop a better understanding of exergame enjoyment. A better understanding of exergame enjoyment can be used to both improve exergames and to help users find exergames that they enjoy. This may lead to more exercise and less health issues for physically inactive people. For this purpose, a questionnaire should be designed and used to assess enjoyment of physical activity in exergames.

A questionnaire is defined as, "A written set of questions that are given to people in order to collect facts or opinions about something" (merriam-webster.com). Currently, there are no existing questionnaires that specifically target exergames, therefore, gaming and physical activity questionnaires that already exist can be observed for inspiration. "Gaming Engagement Questionnaire (GEQ) and Immersive Experience Questionnaire (IEQ), are prominent examples of questionnaires set up in a similar way in order to evaluate player experience. ..these two are available publicly…" (Nordin et al., 2014). Furthermore, the Physical Activity Enjoyment Scale (PACES) is the most commonly used measure of enjoyment of physical activity (Mullen et al. 2011). An exergame enjoyment questionnaire should combine questions from the GEQ, IEQ, and PACES questionnaires because they are reputable and publicly available.

The goal of this IQP is to design and validate a questionnaire for exergame enjoyment. This was accomplished by completing the following steps:

1. Analyzing the current game questionnaires in order to develop a new questionnaire that measures exergame enjoyment
2. Giving the initial rough draft of the questionnaire to multiple focus groups in order to revise it and improve the testing procedure
3. Giving the final draft of the questionnaire to a larger number of users in order to determine how useful it is and how accurate the answers are
4. Evaluate the games used during the 3rd step using the final version of questionnaire

After completing these steps, the data could be analyzed to determine the validity of the questionnaire.

Considering no exergame enjoyment questionnaire currently exists, the development of a questionnaire that assesses exergame enjoyment can be used to influence future exergames. Developers can use the results from this project to improve exergames, such that users experience higher levels of engagement, immersion, and enjoyment. Innovations in exergames can then be used to increase levels of physical activity in young adults, further reducing sedentary lifestyles.

Chapter 2 Background and Exergame Literature Review

This chapter contains background information necessary for this project. Research includes an investigation of existing questionnaires, the components that contribute to exergame enjoyment, considerations for developing an appropriate questionnaire, as well as existing exergames.

2.1 Questionnaires

There are many game questionnaires that exist, and measure different aspects of games. This section reviews several game questionnaires: the Game Engagement Questionnaire (GEQ), the Immersive Experience Questionnaire (IEQ), and the Physical Activity Enjoyment Scale (PACES). Past studies involving the GEQ and IEQ were critiqued to determine the validity and reliability of both questionnaires. Other questionnaires were evaluated through the study *Too Many Questionnaires* (Nordin et al., 2009) to determine important components to assess in an exergame enjoyment questionnaire.

2.1.1 Game Engagement Questionnaire

The Game Experience/Engagement Questionnaire (GEQ) is one of the most well-known questionnaires regarding measuring the enjoyment of digital games (Nordin et al., 2009). It is a general questionnaire with the goal of being able to be applied to any game regardless of genre or mechanics in order to measure their experience. The GEQ has a modular structure and consists of three modules: The Core Questionnaire, The Social Game Module, and The Post-Game Module. Each of these modules are meant to be administered immediately after the game session is finished in the listed order. The GEQ is a non-standardized questionnaire.

The first module, The Core Questionnaire, measures different components such as: Immersion, Flow/Competence, Presence, Absorption, Positive Affect, Negative Affect, and Tension. These components are defined in Figure 2.1.1A.

Figure 2.1.1A Components of The Core Questionnaire

Component	**Definition**
Immersion	"Engaged in the game-playing experience while retaining some awareness of one's surroundings" (Brockmyer et al., 2009)
Flow/Competence	"The feelings of enjoyment that occur when a balance between skill and challenge is achieved" (Brockmyer et al., 2009)
Presence	"Having the experience of being inside a virtual environment" (Brockmyer et al., 2009)
Absorption	"Total engagement in the present experience ... an altered state of consciousness" (Brockmyer et al., 2009)
Positive Affect	The gaming experience gives the player a positive mood
Negative Affect	The gaming experience gives the player a negative mood
Tension	"Mental or emotional strain; intense, suppressed suspense, anxiety, or excitement" (dictionary.com)

Each component has about five corresponding questions. The second module, The Social Presence Module focuses on investigating a user's involvement with other entities, including virtual entities, or mediated entities. (Ex. In game characters, Other human players) The GEQ does state that The Social Presence Module should only be administered if there is a social entity to interact with. The third module, The Post Game Module measures the player's feelings after the game has ended. The purpose of this third module is to assist "naturalistic gaming" (IJsselsteijn, 2016), referring to games that players themselves choose to play.

The number of questions given within the GEQ can vary, however it always contains three modules. The Core Questionnaire, in particular, is longer than the other two modules. Every question within the GEQ is answered in similar fashion, using a five point Likert scale.

2.1.1.1 Game Engagement Questionnaire Critique

The GEQ's ability to accurately assess engagement was validated through a social psychology experiment (Brockmyer et al., 2009). In the study, "The development of the Game Engagement Questionnaire: A measure of engagement in video game-playing" Brockmyer et al. investigated engagement in 107 university-aged males regarding violent video games. The questionnaire was validated through measures of user engagement in these games.

The GEQ that Brockmyer et al. created consisted of 19 questions that specifically measured four components of engagement: Absorption, Flow, Presence, and Immersion (these terms are defined in *Figure 2.1.1.1A*). Each question was ranked by its agreement difficulty. A majority of the negatively phrased questions were classified as the most difficult to agree with, whereas the positively worded questions were deemed the least difficult. Figure 2.1.1.1B provides this agreement difficulty classification, the engagement level for each question, and a respondent distribution of measured engagement. The agreement difficulty classification is relevant to this project considering it recognizes that wording of questions could influence user responses. The respondent distribution is that of a bell curve and has a high peak at the level slightly below zero, which corresponds to an expected average level of engagement. The responses to the questions (N = No, M = Maybe, Y = Yes) provided an expected distribution of engagement relative to each question's agreement difficulty, or how likely a user was to agree with a given question. This is evidenced by the respondent distributions bell curve structure, considering a majority of the users were deemed to hold an average level of engagement, while few held extreme levels of engagement. It was noted that the scale of agreement difficulty for each question allowed the GEQ to produce a distribution of engagement levels. By asking questions that were considered difficult to agree with, the GEQ was able to measure users with a high potential for the deepest level engagement. This pattern was also noticed with users who expressed the lowest potential for engagement. Therefore, considering the developed GEQ was able to identify a breadth of engagement levels and patterns, the questionnaire was deemed to provide a "Psychometrically strong measure of the type of engagement elicited while playing video games" (Brockmyer et al., 2009).

The developed GEQ was considered to accurately measure violent video game engagement; however it had limitations that affected its accuracy. In the respondent distribution, few respondents fell outside of one standard deviation from the mean level of engagement. This indicates that the developed GEQ should, "Include more items at the higher level of engagement … and eliminate lower engagement items" (Brockmyer et al., 2009). As a result, the developed GEQ would be improved by including more questions that assess extreme levels of engagement. This would produce a distribution where more respondents fall further than one standard deviation from the mean; thereby assessing users that express extremely low or high levels of engagement. Furthermore, it was noted that it should have been determined "how the GEQ performs with ... diverse ethnic and age groups, as well as with children and adults" (Brockmyer et al., 2009). The concern of diversity was mentioned considering the test users were all university-aged males. The GEQ may perform differently when administered on a more diverse group of users. Additionally, component overlap was a concern. Questions sixteen and seventeen were noted to potentially cause "overfitting or item redundancy" (Brockmyer et al., 2009). This indicates that there were too many similar questions, or that some questions were oversampling a particular component.

This study emphasizes several major aspects that are important to this project regarding the creation of the questionnaire. The following qualities are essential:

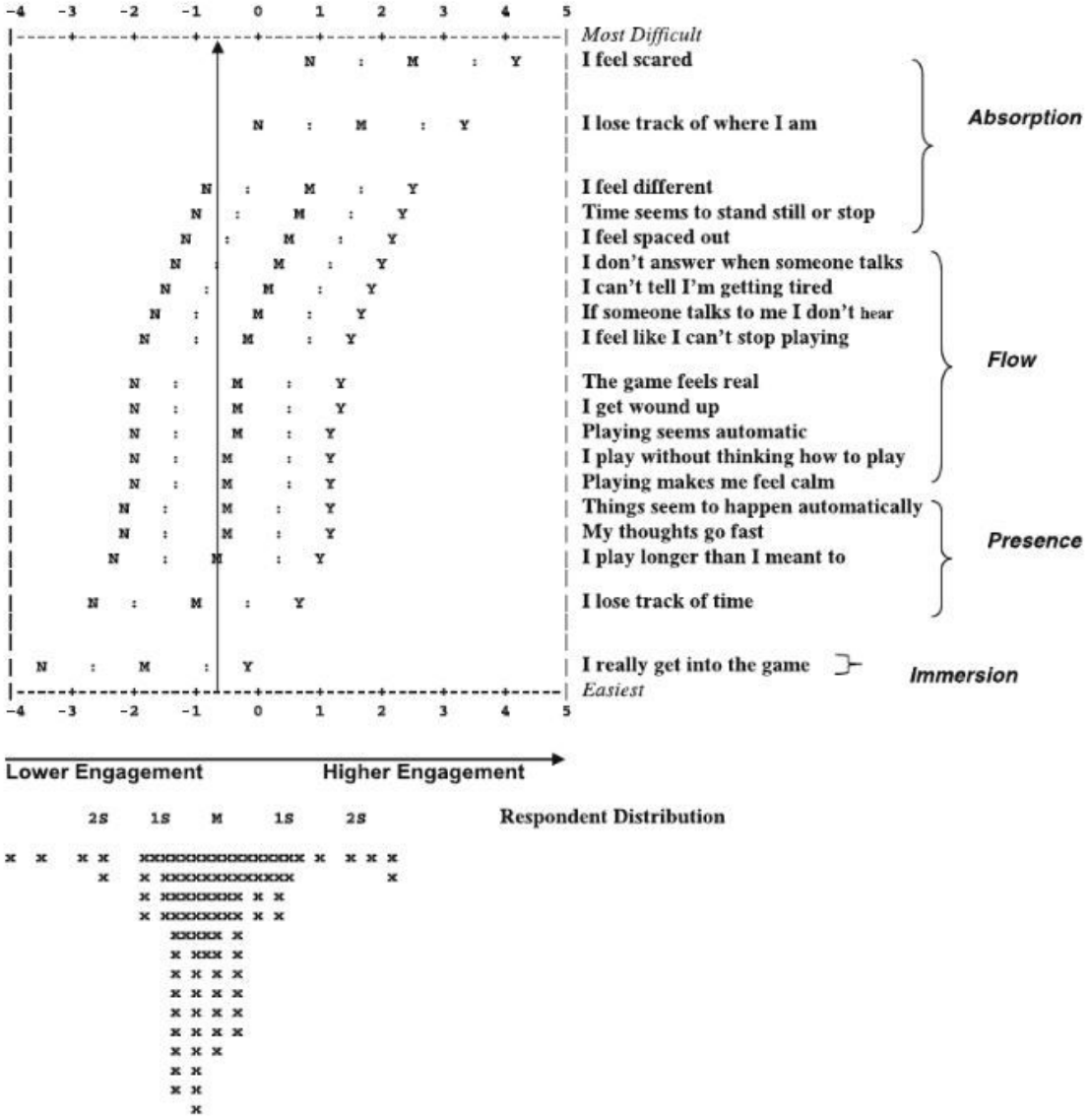
- Extreme questions: required to produce a respondent distribution where users are not all grouped at the mean level of engagement. It is important to be able to assess those users who experience extremely low or high levels of engagement.
- User diversity: the questionnaire may have performed differently across various age groups and ethnicities. For that reason, it is important to test a variety of participants.
- Question agreement difficulty: the questionnaire should have questions that vary in agreement difficulty. Essentially, the questionnaire should have a range of questions that are considered easy to agree with to difficult to agree with, including extremes of each.

Figure 2.1.1.1A Example Questions from the GEQ (Brockmyer et al., 2009)

Table 1
Game Engagement Questionnaire (GEQ) items.

1	I lose track of time
2	Things seem to happen automatically
3	I feel different
4	I feel scared
5	The game feels real
6	If someone talks to me, I don't hear them
7	I get wound up
8	Time seems to kind of stand still or stop
9	I feel spaced out
10	I don't answer when someone talks to me
11	I can't tell that I'm getting tired
12	Playing seems automatic
13	My thoughts go fast
14	I lose track of where I am
15	I play without thinking about how to play
16	Playing makes me feel calm
17	I play longer than I meant to
18	I really get into the game
19	I feel like I just can't stop playing

Figure 2.1.1.1B Respondent data and engagement levels (Brockmyer et al., 2009)



2.1.2 Immersive Experience Questionnaire

The Immersive Experience Questionnaire (IEQ) is another well-known and commonly used questionnaire (Nordin et al., 2009). Similar to the Game Experience Questionnaire, the Immersive Experience Questionnaire is designed to be applicable to all games and tested as such. The IEQ focuses specifically on the Immersion aspect of games.

It measures the Immersive experience of a game, categorizing immersion into three levels: Engagement, Engrossment, and Total Immersion. Engagement normally occurs after the user has invested time and effort into the game, once the user has "Overcome the Barrier of gamer preference" (Oum, 2009). Engrossment refers to when the gamer's emotions are able to be affected by in game events. Total Immersion is the last stage, referring to when the user is cut off from reality, focusing entirely on the game.

The IEQ measures the user's response to the game, allowing it to provide insight on how immersed the user is into the game. It shows the user's emotional and cognitive involvement, and dissociation from the real world while playing the game.

The IEQ uses the same five point Likert scale that the GEQ uses. However, the IEQ specifies positivity and negativity with its questions, one that strongly disagrees and five that strongly agree. The IEQ, like the GEQ, is also meant to be administered after the user plays the game rather than before.

2.1.2.1 Immersive Experience Questionnaire Critique

The IEQ's ability to accurately measure immersion was validated through a Human-Computer Interaction study. The study, "Measuring and Defining the Experience of Immersion in Games" (Cox et al., 2008) consists of three experiments that attempt to determine if immersion can be measured quantitatively. Experiment one investigated the effects on immersion when the user switched from an immersive to a non-immersive task. Experiment two studied user eye movement during immersive tasks. The last experiment, experiment three, investigated the effect of user interface interaction speed during immersive tasks. The overall goal of the study was to develop a clear meaning of the term, "Immersion" such that a better understanding of how users engage with video games can be reached.

The first two experiments used the same IEQ, while the third experiment used an entirely different IEQ. The first questionnaire consisted of 33 questions in total, the second was 31 questions. Each questionnaire had five target components consisting of: cognitive absorption, emotional involvement, real world dissociation, challenge, and control. Users were to respond to each question in the first questionnaire with a Likert Scale in which one corresponded to strongly disagree and five corresponded to strongly agree. The questionnaire for experiment three used a similar Likert Scale for user responses; "Not at all" was lowest on the scale, and "A lot/Very much so/Very Aware" were highest on the scale. The questionnaire for the first two experiments targeted the user themselves, where nearly every question was phrased reflexively. Experiment three's questionnaire asked questions in such a way that they were targeted towards the game itself, rather than the personal feelings of the user. Lastly, the immersion scores were computed by summing user responses to the questions.

Experiments one and two were deemed to be successful in quantitatively measuring immersion. Experiment two was similar to Experiment one, which "Would appear that the immersion questionnaire developed was a successful indicator of immersion" (Cox et al., 2008). This conclusion was reached considering the levels of user immersion were higher than those of the control tests in both experiments. Experiment three found that "one can suggest that emotional involvement appears to be a key factor in immersion" (Cox et al., 2008). This suggests that for a given IEQ to produce reliable measures of immersion, the questionnaire should assess

the user's emotional involvement. All three experiments' questionnaires were found to accurately produce quantitative measurements for immersion regarding the five target components. The experiments concluded overall that the two immersive experience questionnaires developed were valid in obtaining quantitative measures of immersion.

The developed questionnaires were able to accurately assess their target components yet could be further refined. It was noted that the researchers of this study, "plan to develop their immersion questionnaire further through the use of qualitative measures, by interviewing gamers" (Cox et al., 2008). By assessing gamer thoughts before creating the questionnaires, through the use of focus groups, the questionnaires could have resulted in more accurate measures of immersion. Similar to the GEQ in section 2.1.1.1, diversity of users was mentioned as a concern for the same performance reasons.

This study is relevant to this project regarding the creation of the questionnaire. The results of this study emphasize the importance of assessing user emotional involvement, focus groups, and the effects of diversity in users. By asking questions about the user's emotional involvement in the questionnaire, it is possible to measure extreme levels of immersion. The study also mentioned that utilizing user interviews, essentially focus groups, could have improved their questionnaire. This is important for this project considering focus groups can be used as a method of validating the accuracy of the questionnaire. Lastly, diversity was a concern similar to the study in section 2.1.1.1, such that the questionnaire may perform differently when administered on other user age and ethnicity distributions.

2.1.3 Physical Activity Enjoyment Scale Questionnaire

The *Physical Activity Enjoyment Scale*, also known as *PACES*, is a questionnaire that examines the feelings of users pertaining to physical activity (Mullen et al., 2011). There are multiple variants of the *PACES* questionnaire, but they generally all consist of a list of statements regarding physical activity. Each statement describes a type of opinion or feeling that the user could have. With a 5-point Likert-type scale, users are given the opportunity to show how much they agree or relate with each of the 15-20 statements. The *PACES* questionnaire is applicable to people of all ages and has been used to test both children and adults.

This paper-and-pencil-administered questionnaire begins with the stem, "When I am physically active…" and proceeds to list various statements that serve as possible responses for the user such as "I enjoy it", "It makes me sad", or "It gives me a feeling of success" (Mullen et al., 2011). The variety of statements not only determines if the user enjoys physical activity, but it analyzes *why* the user enjoys or does not enjoy exercising.

In order to understand the appeal of exergames, it is essential to understand if users enjoy exercise because it feels good, because of the feeling of success that comes from doing it, or for some other reason. The study, "Measuring enjoyment of physical activity in older adults: invariance of the physical activity enjoyment scale (paces) across groups and time" supports the validity of the PACES questionnaire as a measure of enjoyment for older adults, regardless of the type of exercise (Mullen et al., 2011). Furthermore, the study, "Measuring Enjoyment of Physical Activity in Children: Validation of the Physical Activity Enjoyment Scale" also supports the validity of the PACES questionnaire as a measure of enjoyment for children (*Journal of applied sport psychology*, 2009). There is minimal information about the PACES questionnaire being used for college students or young adults. However, the *PACES* questionnaire may be a useful precedent for writing physical activity-related questions, because it is valid for other age groups. It helps complete the exergame enjoyment questionnaire by including the exercise component that the gaming questionnaires ignore.

2.1.4 Other Questionnaires

Too Many Questionnaires: Measuring Player Experience Whilst Playing Digital Games, is a study that analyzes and compares a few prominent game questionnaires in order to position them in relation to each other (Nordin et al., 2009). In the paper, they list the most prominent questionnaires that measure the engagement for digital games as well as investigate how current questionnaires could be improved. The conclusion on the paper is that further investigation is needed to created better quality questionnaires to reduce confusion.

The article gives a comprehensive list of questionnaires that measure the engagement for digital games, and gives more analysis and focus on the two most prominent questionnaires, the GEQ, and the IEQ. It contrasts the two questionnaires and analyzes the differences between the two. The paper hypothesizes that scores between the two questionnaires are correlated. Additionally, for each of the questionnaires, all the measured components are listed and shown, revealing much overlap between the different questionnaires (Figure 2.1.4). The categories of 'overlap' are described in Section 2.2.2.

There are many questionnaires listed other than the GEQ and the IEQ. The list of questionnaires include Flow Questionnaire, Presence Questionnaire, Gameflow Questionnaire, Player Experience of Needs Satisfaction (PENS), and Social Presence in Gaming Questionnaire (SPGQ) (Nordin et al., 2009). This project based many of the questions off of the GEQ and IEQ, however, the other questionnaires should be considered in future work.

The Player Experience of Needs Satisfaction (PENS) has a criteria for what players need for a game to be successful. A few examples of this are needs for competence, freedom, and relatedness. These categories are very similar to the components measured in other questionnaires. PENS is a questionnaire applied to entire genres of games such as first person shooter or simulation games.

The Social Presence in Gaming Questionnaire (SPGQ) focuses on the interlayer interaction in games. This often takes the form of interacting with other, real, players.

The Flow Questionnaire, Presence Questionnaire and Gameflow Questionnaire, are specific questionnaires that target very specific elements of games, and are used much less than the other four questionnaires.

This project focused more on the GEQ and IEQ for a few specific reasons. GEQ and IEQ are the most prominent questionnaires, and are still used today, representing the industry standard. While PENS and SPGQ are still used, their usage is too a much lesser extent. PENS

specializes in defining genres and giving an overview for what elements need to be in a game and if they are there, and the SPGQ is specifically designed to measure the social element in games. They measure very specific aspects of games, and thus cannot be as widely implemented as the GEQ and IEQ.

Figure 2.1.4 From *Too Many Questionnaires* Paper (Nordin et al., 2009)

Questionnaire	**Components**
GameFlow Questionnaire	Concentration A sense of challenge Player skills Control Clear goals Feedback Social interaction Immersion
Game Engagement Questionnaire (GEQ)	Absorption Flow Presence Immersion
Player Experience of Needs Satisfaction (PENS)	Competence Autonomy Relatedness Presence (Immersion)
Social Presence in Gaming Questionnaire (SPGQ)	Psychological involvement (empathy) Psychological involvement (negative feelings) Behavioural engagement
Flow Questionnaire	Clear goals High concentration Reduced self-consciousness Distorted sense of time Direct and immediate feedback Balance between ability level and challenge A sense of personal control Intrinsically rewarding activity
Presence Questionnaire	Control factor Sensory factor Distraction Realism factor
Immersive Experience Questionnaire (IEQ)	Emotional involvement Cognitive involvement Real world dissociation Challenge Control

2.2 Understanding Exergame Enjoyment

This section elaborates on how enjoyment is expressed and how our proposed questionnaire could target components of enjoyment. It is necessary to define "Exergame Enjoyment" and also to understand how participants express enjoyment in order for the questionnaire to accurately measure exergame enjoyment.

2.2.1 Defining Exergame Enjoyment

In order to assess user enjoyment of exergames, it is essential to first define the term "exergame enjoyment". An exergame is defined as, "A video game that also provides the player with physical exercise" (yourdictionary.com). Enjoyment is defined as, "A feeling of pleasure caused by doing or experiencing something you like" (merriam-webster.com). From these two definitions, "exergame enjoyment" can be defined as, "A feeling of pleasure caused by playing a videogame which provides physical exercise". The components contributing to this "feeling of pleasure" will be considered in subsequent sections.

2.2.2 Components to Test

According to Nordin et al., there are many components that supposedly contribute to video game enjoyment. *Figure 2.1.4* shows key components of video game enjoyment:

Immersion	● Absorption ● Behavioral engagement ● Cognitive involvement ● Concentration ● Distorted sense of time ● Distraction ● Emotional involvement	● Flow ● Immersion ● Presence ● Psychological involvement ● Real world dissociation ● Reduced self-consciousness ● Sensory factor
Intrinsic Reward	● Challenge ● Competence ● Feedback	● Intrinsically rewarding ● Player skills
Control	● A sense of personal control ● Autonomy	● Control
Miscellaneous	● Clear goals ● Realism factor	● Relatedness ● Social interaction

The components were categorized based on broad categories. The first component, **immersion**, is defined as, "the state of being deeply engaged or involved; absorption" (dictionary.com). The second category of components relate to **intrinsically rewarding activity**. An intrinsic reward is defined as, "an outcome that gives an individual person satisfaction such as that derived from a job well done." (businessdictionary.com). Next, the third set of components relate to **control**, which is defined as, "the power to influence or direct people's behavior or the course of events." (oxforddictionaries.com). These three components were considered for inclusion to assess the game element of the proposed EEQ. *Table 5.4A* shows questions that fall into these categories as well as an additional physical activity category.

The miscellaneous components category were ignored for the purpose of this study. The "Relatedness" and "Social interaction" items both concern the social or multiplayer aspects of video games. "Clear goals" and the "Realism factor" do not seem like necessary qualities of an enjoyable exergame. "Clear goals" is not as relevant as the other categories considering the test subject only played the game for ~30 minutes. Furthermore, a game does not necessarily need to be 'realistic' to be enjoyable.

2.3 EEQ Considerations

This chapter contains ideas that should be taken into account when creating a questionnaire. The sections focus on 'conducting' - how to conduct questionnaires appropriately, 'budgeting' - choosing a reasonable number of questions, and other basic mistakes that need to be avoided when writing the questions.

2.3.1 Conducting and Budgeting

Conducting and budgeting for a survey can impact the results gained and conclusions gleaned, so it is very important for one to know all the factors that go into each decision (FluidSurveys, 2016).

For each questionnaire, the number of questions that should be asked will vary. While it is important to be thorough with questioning, making sure one has enough information to work with, it is incredibly important to make sure that the EEQ is not too long. While many papers and articles have their own numbers, all are within the same range, that surveys should not take more than 10-15 minutes to complete. Each question averaging 5-20 seconds to complete.

Each "component" in a questionnaire should have five elements measuring it to make sure the date is accurate. For example, if you were to test five components with five components each, that would yield a total of approximately 25 questions.

For conducting surveys there are two main ways to do so, first would be traditional interviews where a proctor speaks to the subject answering the questionnaire in some way. The second method is online, where there is no proctor involved. Each approach has its own advantages and disadvantages. An online survey can garner a large amount of responses, giving more data to work with as well as reaching a wider and more diverse audience. However, traditional surveying offers much more control with the presence of the proctor or interviewer. For new surveys and questionnaires, a controlled environment is needed to make sure the data is being gathered in an unbiased and fair manner. Also, traditional surveying is more simple to change. As such, it was decided that the EEQ would first be tested with smaller, and proctored groups, before releasing the questionnaire to the internet.

2.3.2 How to Ask Questions and Mistakes to Avoid

It is important to ask questions in a professional manner. The points below explain how to ask questions correctly (Beadell, 2016).

1. **Use simple, direct language.**

   The questions should be short, clear, and with simple wording.

2. **Be specific.**

   Some concepts might be different for different people. For example, in the question "do you often go to the gym", the word "often" could mean once per week for one person but five times per week for another person. Instead, the question should be worded as "on average, how many days do you go to the gym each week?"

3. **Break down big ideas into multiple questions.**

   If the object of one question is multifaceted, it is better to break the question into multiple smaller questions that assess different aspects of the overall idea. For example, instead of asking "customer satisfaction", asking customers opinions on statements such as if they enjoy using the product, if the product meets their need, and if they will purchase for this company again.

4. **Avoid leading questions.**

   Sometimes a researcher's opinion will influence how the question is asked, so that it can subtly encourage the respondents to answer the question in a certain way. For example, asking "Do you think the school should cut the gym budget to pay crossing guards?"(Beadell, 2016) would likely prompt a different answer than asking, "Should the school employ crossing guards to protect our children?"(Beadell, 2016) even though both questions are related to the same topic.

5. **Assess one component per question (avoid loaded and double-barreled questions).**

   Each question should ask one thing and one thing only. A double-barreled question such as, "do you eat fruits and veggies on a daily basis?" (Beadell, 2016) would be tricky to answer for people who only eat fruits or veggies but not both.

6. **Use interval questions and provide enough options.**

   Interval questions provides a much wider variety of answers than Yes/No and multiple choice question. Yes/No or multiple choice questions could easily block possible true answers, while asking interval questions, respondents are more able to express their opinion using a Likert Scale consisting of the possible responses, "Strongly Disagree, Disagree, Neither Agree nor Disagree, Agree, or Strongly Agree." E.g. If a question asked participants how they felt about the statement 'I frequently eat ice cream', then they could agree or disagree, they could strongly agree or strongly disagree, or they could choose 'neutral' if they are on the fence.

7. **Avoid bias.**

   Bias could happen on a small scale. Asking "How important do you think content marketing is?" (Beadell, 2016) followed by "How much do you plan to invest in content marketing next year?" could very well lead to bias if someone says yes to the first question because he/she would probably inflate the dollar amount for the second question. The solution is to randomize the order of the questions.

2.4 Exergame Research

A list of exergames was needed for this project in order to determine possible exergame choices. Hope Wallace researched, classified, played, and then ranked nearly 80 exergames available on both Android and iOS mobile-phone operating systems (Wallace, 2016). After evaluating each of the games, an elimination process was used to determine fifteen, "Best Exergames". Figure 2.4A describes the elimination process used.

The elimination process first discarded applications that were not free-to-play, this is because this project was intended to target free-to-play games. Games that could not be opened or had bugs that made gameplay difficult were also eliminated. Furthermore, applications that could not be defined as games and had no exercise elements were discarded. Lastly, unpopular games (those with few ratings and reviews) and those that were considered to only be step-counter applications were not considered. After this process was applied, a list of fifteen exergames remained (shown in Figure 2.4B).

Figure 2.4A Exergame elimination process

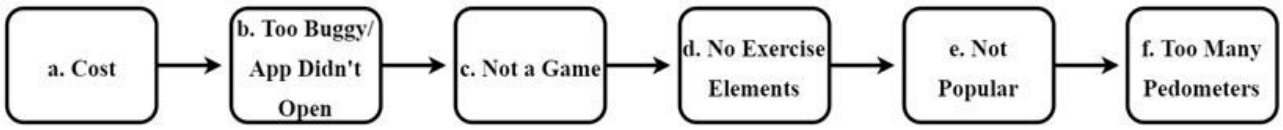


Figure 2.4B Top fifteen exergames (in no particular order)

Game	**Description**
Fitbit	Distance/Step counter, uses “Fitbit” wearable pedometer.
Fitocracy	Personal workout trainer, achievements for workouts, uses social media.
Geocaching	Users walk to locations on the real world map to explore Geocaches.
Ingress	Users walk to locations on the real world map to explore landmarks.
Just Dance Now!	Users dance to music with on screen dance instructions.
Nike+ Running	Distance/step counter with achievements and goals.
Pokémon GO	Users walk to locations on the real world map to catch Pokémon.
Resources Game	Users walk around the real world map to collect “resources”.
S Health	Personal workout trainer, achievements and goals for workouts.
Shape Up Battle Run	Racing game that requires user to walk, jog, or run with achievements and goals.
SpekTrek Light	Users walk or run on the real world map and catch in-game “ghosts”.
TableZombies Augmented Reality	Augmented reality game that requires user to walk in order to move around within the game.
Turf Wars	Users control territories on the real world map and walk or run based on in-game events such as evading the police.
Walkr: Fitness Space Adventure	Users walk to explore the in-game map and earn achievements and unlocks.
Zombies, Run!	Walking/running game with challenges and achievements.

Chapter 3 Methodology

This chapter describes the methods by which this project was conducted and how the exergame enjoyment questionnaire was validated. Specifically, the project was completed by following these steps:

1. Performed background research on the following topics:
    a. Gaming and physical activity questionnaires
    b. Exergame enjoyment
    c. Considerations for developing a reliable questionnaire
    d. Existing exergames

    The background research is presented in Chapter 2 of this report.
2. Developed a rough draft for the Exergame Enjoyment Questionnaire (EEQ) and gave it to a small focus group after they had played one of the two exergames: *Pokémon GO or Just Dance Now*. The focus group was used to validate the draft EEQ. The focus group validation process, participant demographic, and exergame testing procedures are presented in the sections that follow. The questionnaire components are also provided, along with the method for collecting participants' responses and the system for scoring the EEQ responses.
3. Finalized the EEQ and gave it to a much larger group of users after they also played one of the same two exergames.
4. Scored and analyzed the results of the testing from Step 3.

3.1 Validation of EEQ using Focus Groups

Focus groups were conducted in order to validate that the questionnaire accurately assessed exergame enjoyment. First, the procedure for playing each exergame was conducted with a small number of users. Afterwards, the participants filled out the proposed EEQ. Lastly, a discussion on exergame enjoyment followed where the users were asked open-ended questions about their experience. The users' thoughts, comments, questions, and reactions to the exergame were recorded during this conversation. The user response data was then compiled and coded in order to develop categories of responses. These categories were used to find relationships between the developed EEQ and the user responses. This iterative process was used to improve the EEQ before deployment and experimentation with a large user group.

3.2 Participant Distribution

The participants involved in this project were all Worcester Polytechnic Institute (WPI) students. With the WPI student gender distribution consisting of roughly two-thirds male and one-third female, it was expected that the gender distribution for this project would roughly follow that of WPI. Figure 3.2A displays the gender distribution for this project which was approximately 58 percent female and 42 percent male. A majority of the participants were within the 18 - 22 age range, however there was one participant aged 27. Figure 3.2B provides this project's participant age distribution.

Additionally, relevant information regarding exercise and exergame knowledge was obtained in this project. Figure 3.2C displays the distribution of participant exergame knowledge prior to engaging in the study. Fifteen participants said that they knew what an exergame was, while 23 did not. Lastly, each participant was asked how many hours they exercise in a given week. Figure 3.2D shows the distribution of participant exercise. This distribution shows that the participants involved in this study engaged in at least some amount of exercise in a given week.

Figure 3.2A: Participant gender distribution.

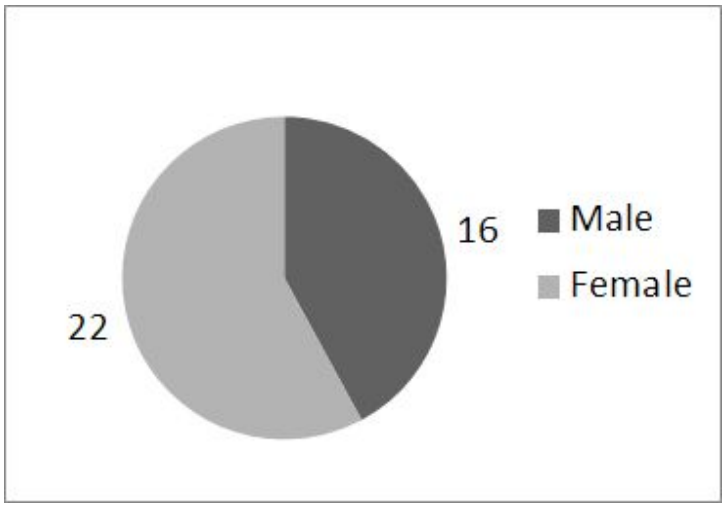


Figure 3.2B: Participant age distribution.

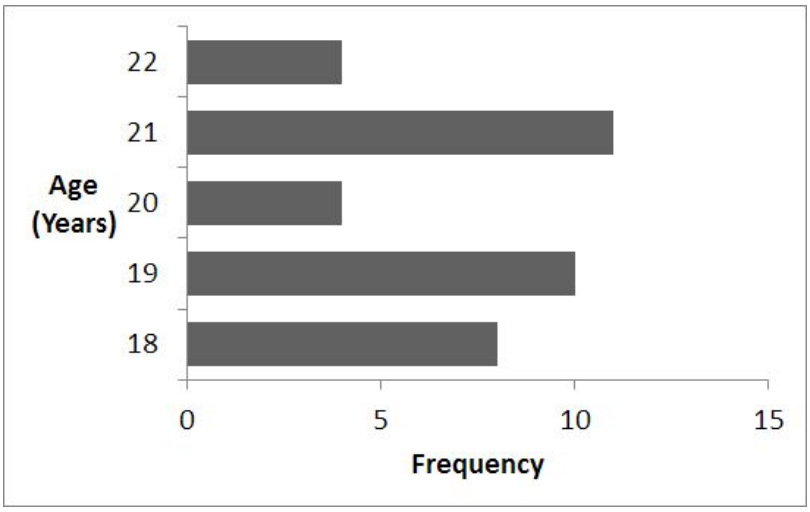


Figure 3.2C: Participants' knowledge of what exergames are.

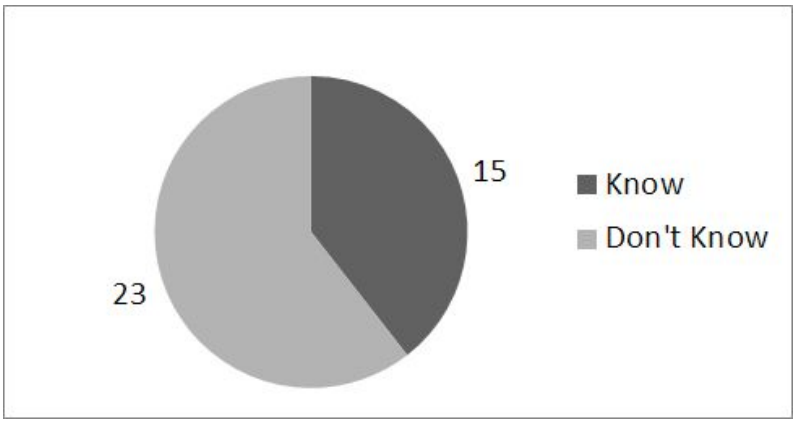
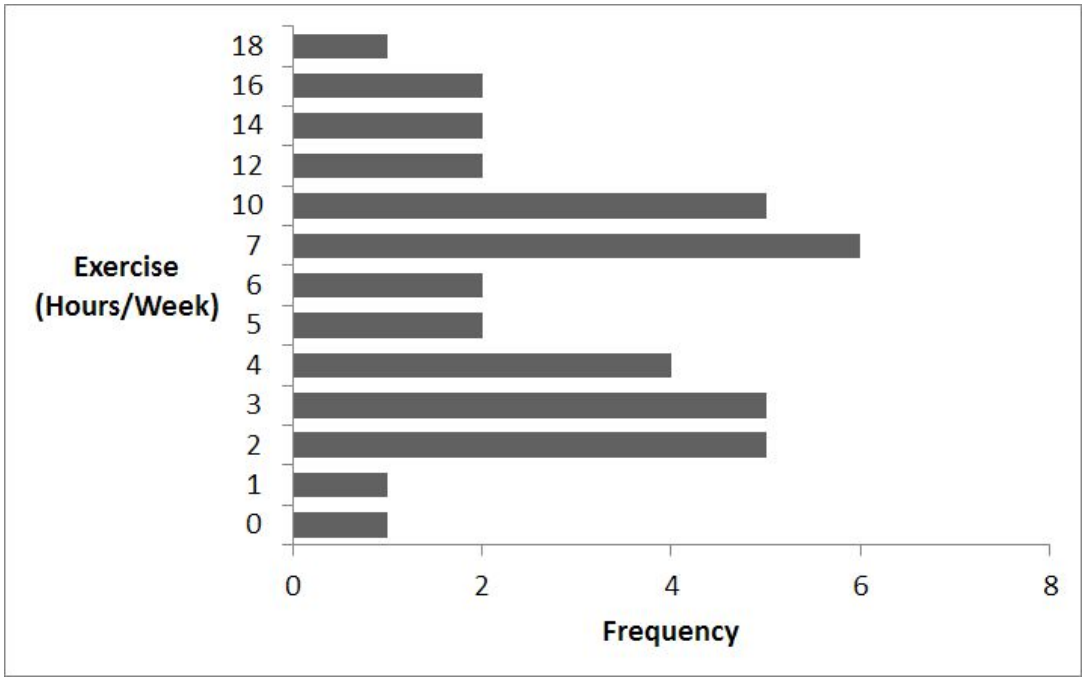


Figure 3.2D: Exercise Performed by Participants Per Week



3.3 Exergame Procedures

The two exergames selected for testing were *Pokémon GO* and *Just Dance Now*. Participants played the games for approximately thirty minutes. Within that time, they were given specific objectives for each game in order to reach a minimum amount of participation. For *Pokémon GO*, participants were to walk around campus while playing the game. They were instructed to reach at least ten Pokéstops; all of which were located on the WPI campus. They could choose to keeping moving around to new locations or stand still for the remainder of the testing period. *Just Dance Now!* users were to play in a classroom with the game presented on a projector screen (Image 3.3A). Users played as many songs as possible within the 30 minutes timeframe.

Image 3.3A: Participants playing Just Dance Now!

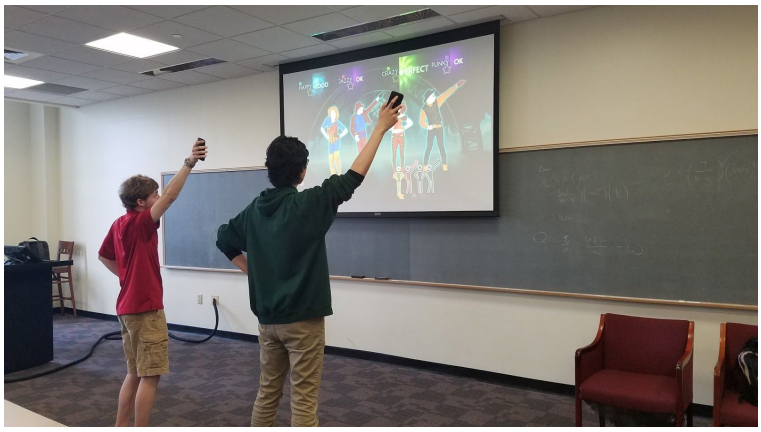


Image 3.3B: Participants being administered the Exergame Enjoyment Questionnaire

Image 3.3C: Participants and researcher during a focus group discussion

3.4 Questionnaires

Each test subject took a pre-questionnaire before playing the exergame. The questionnaire had two components. A Pre-Exergame Questionnaire was given prior to the user playing the game. The Exercise Enjoyment Questionnaire was administered immediately after 30 minutes of gameplay. Refer to Chapter 5 for more information regarding the EEQ components.

The Pre-Exergame Questionnaire, shown and explained in *Table 5.5*, was used to determine the user's previous experience of playing video games and exercising by asking five short-answer type questions. If there was a relationship between exergame enjoyment and prior experience, the Pre-Exergame Questionnaire helped display that relationship if test subjects were categorized by their answers.

The Exergame Enjoyment Questionnaire was the main focus of this project. It was conducted in a setting that would allow the users to focus on the questions without external influence. The questionnaire was given on paper, with 20 questions split between two pages (Image 3.3B). For both exergames, users answered the questionnaire in a quiet classroom setting. A researcher was present in the classroom, but not watching the user or influencing the user's responses. The researcher did not answer user questions related to the interpretation of questions on the EEQ. If a user asked the researcher to explain a question, the researcher's response was always, "Answer the question to the best of your ability."

3.5 Scoring the EEQ

A user's enjoyment of the exergame was scored based on their answers to the questions. Each answer on the EEQ was given as a five choice Likert scale, ranging from "Strongly Disagree" to "Strongly Agree". If a statement was phrased positively, then the possible answer choices would add points to the score as shown in Figure 3.6A. E.g. "The exercise in this game made me feel good." In contrast to a statement that is phrased negatively, the possible answer choices would add points to the score in the reverse order, as shown in Figure 3.6B. E.g. "I did not feel like I wanted to keep playing." The scores for each question were summed to yield a total enjoyment score for each user. A higher overall score (maximum of 5) indicates that the user enjoyed the game more than another user with a lower point total (minimum of 1).

Figure 3.6A: Point Values for a Complimentary Statement

Answer Choice	**Point Value**
Strongly Agree	5
Agree	4
Neutral	3
Disagree	2
Strongly Disagree	1

Figure 3.6B: Point Values for a Negative Statement

Answer Choice	**Point Value**
Strongly Agree	1
Agree	2
Neutral	3
Disagree	4
Strongly Disagree	5

Chapter 4 Exergame Choices

This chapter explains the process of selecting two exergames to test in this project. In order to choose two exergames from a list of fifteen, the properties and characteristics of each exergame were compared. Characteristics that were considered extremely important for an exergame to be tested for this project are: gameplay difficulty, testing difficulty, game elements, and physical activity elements. Additional characteristics that were considered less important but still factored into the selection of exergames were the types of physical activity provided and game popularity.

4.1 Game Choice Criteria

This section describes desired qualities of exergames to be tested for this project and why they were desired. The goal was to select two exergames that fulfilled these desired qualities and that also noticeably differed from one another. The six characteristics that were considered significant factors in selecting the two exergames to test were mentioned above: the game element, the exercise element, the testing difficulty, the gameplay difficulty, the type of exercise provided, as well as the game popularity.

Game Aspects: Exergames should incorporate physical activity with game elements, such that they are not just workout trainers or simple exercise applications. Some examples of game elements are: fantasy, goals, and competition, to name just a few. If the application allows the player to explore an unreal world or to pretend they are somebody else, as they are exercising, then the application might be considered a game. *Zombies, Run!* players get to enjoy this fantasy element by pretending to be apocalypse survivors in a world full of zombies. If the application provides the player with some kind of interesting goal, then it might also be considered a game. *Pokémon GO* players are exercising, but they are distracted by a goal to discover and catch Pokémon. Finally, if an application allows players to compete with others, then it might have game elements. *Just Dance Now* players are focused on earning higher scores by dancing better than their friends and opponents. These game elements provide the player with an entertaining distraction from the physical activity that is taking place.

Exercise Aspects: An exergame should also include exercise elements. This could be any form of physical exertion that benefits the user's health in some way. The fifteen considered exergames involve walking, running, or dancing - all of which are activities that involve considerable movement. If the application did not require any exercise, then it would just be a regular mobile game, but not an exergame.

Testing Difficulty: 'Testing difficulty' is another important factor for this project. In order to develop a stronger understanding of exergame enjoyment, a large sample size of people should be tested. For that reason, the testing procedure should be convenient and easy to execute if possible. For instance, *TableZombies* is a game that requires various extra pieces of equipment

beyond just the user and mobile device, making it more difficult to test than most of the other exergames on the list.

Gameplay Difficulty: On a similar note, gameplay difficulty is also a detail that needs to be considered. Users need to be taught how to play the exergame before they can even participate in the testing. If a particular exergame has an extensive learning curve, then it would require the proctors more time to explain the game to the participants. Furthermore, a complicated game could result in confusion for the players - and if players are having trouble understanding the exergame, then their questionnaire answers may be an unreliable review of the game. The beauty of a game like *Just Dance Now* is that the premise is extremely simple; follow the actions of the dancing figure on screen. A simple game can be tested within an hour-long block of time and also allows for more trustable data.

Exercise Types: Another important quality to consider for exergame selection is the type of exercise that the different games offer. All forms of exercise will benefit the user's health, however, there is a limited variety among the fifteen exergames. Most of the games involve walking or running whereas *Just Dance Now* is the only game that stands out by providing something else: dance. The issue with choosing two games that incorporate the same form of exercise is that the testing results would give an incomplete picture of exergame enjoyment. There are many ways to exercise and understanding why a user enjoys walking games would not be sufficient for understanding why a user enjoys exergames in general. A different type of game allows the concept of 'exergame enjoyment' to be observed from another angle; more information can be gained.

Popularity: The last criterion is game popularity. The hope is that participants will like the games that they are playing, so that exergame enjoyment can be better understood. For that reason, if a group of games share a similar premise, then the most popular game in the group would be more appropriate for testing. *Geocaching*, *Ingress*, and *Walkr: Fitness Space Adventure* are games that incorporate walking and exploration. However, these games are all significantly less popular than *Pokémon GO*, which provides the exact same type of activity. Furthermore, a more popular game allows players to experience play with other real people.

4.2 Top Fifteen Exergames and Elimination Rationale

Each of the fifteen exergames' (Section 2.4) qualities were further examined. The two most suitable exergames for testing exergame enjoyment were selected based on the criteria described in Section 4.1. The rationales for ignoring/eliminating the other thirteen exergames can be reviewed in Figure 4.2.

In this list, *Pokémon GO* was chosen first because of its simplicity and overall popularity. The goal is to understand exergame enjoyment, which means the result of our questionnaire for *Pokémon GO* should be generally positive. The other game should be different enough from *Pokémon GO* to test the viability of the questionnaire on different exergames. *Just Dance Now* stood out as the other game for testing and thirteen other games were eliminated due to a variety of reasons described in Figure 4.2.

Figure 4.2 Final Exergame elimination rationales (Pokémon GO and Just Dance Now omitted)

Exergame	Elimination Rationale
Fitbit	Requires additional technology: the Fitbit Watch.
Fitocracy	Does not have many game elements. Targeted towards fitness rather than game elements.
Geocaching	Too similar to *Pokémon GO*. Players are required to move far from the WPI campus to play. This study took place on the WPI campus with students.
Ingress	Too similar to *Pokémon GO* and less popular.
Nike+ Running	Does not have many game elements. The game seems to be more targeted towards fitness rather than the game aspect.
Resources	Too similar to *Pokémon GO* and is less popular.
S Health	Does not have many game elements. Targeted towards fitness rather than game elements.
Shape Up Battle Run	Has a good balance between game elements and physical activity, but it is less popular due to its simplicity.
SpekTrek Light	Extremely similar to *Pokémon GO*, but it is less popular.
TableZombies	Requires additional components
Turf Wars	Similar to *Pokémon GO*. Focuses on control of an area instead of catching imaginary creatures. Requires a large user base in order to maximize user experience.
Walkr: Fitness Space Adventure	Extremely similar to *Pokémon GO*, but it is less popular.
Zombies, Run!	Lengthy testing periods that can last as long as 40 minutes. It is a very simple running game with a focus on the storyline.

4.3 Exergames Selected for Evaluation

This section explains the rationales for selecting the two exergames that were tested for this project. The chosen exergames, *Pokémon GO* (Niantic Labs, 2016) and *Just Dance Now* (Ubisoft, 2014) were considered to hold all of the desired characteristics and were also identified as contrasting exergames. This section also describes general gameplay for each game.

4.3.1 Pokémon GO

Pokémon GO (Niantic Labs, 2016) is designed around the world map, where users walk to destinations to collect items, catch *Pokémon*, battle with other players' *Pokémon*, and earn achievements. Typically, players will spend most of their time in the game throwing *Pokéballs* at *Pokémon* monsters that appear on the map in order to catch and collect them.

- **Pokémon:** a variety of fictional creatures that reside in the wild or belong to humans
- **Pokéballs:** small spherical balls that can be thrown at Pokémon in order to capture them
- **Pokéstops:** locations where humans can go to collect various different items

This game was considered to be the best exergame to test in this project. This decision was primarily based off of the game's overwhelming popularity, becoming the "Most downloaded mobile game in its first month" (Swatman, 2016). Additionally, *Pokémon GO* exhibits the criteria mentioned in Section 4.1. *Pokémon GO* was considered to focus heavily on game elements and the "Fun-Factor" rather than just physical activity. Additionally, the game only provides physical activity through walking. These observations were significant for selecting two contrasting exergames.

Pokémon GO also provided a simple user testing procedure. Considering that the game was designed for players to walk in order to play, the procedure for user tests was relatively simple. No setup was required: the user can take the mobile device and immediately start playing. Additionally, *Pokémon GO* was considered easy to learn and play. Interfaces and controls in the game were designed such that their functions and usage would be obvious to users who have never played the game prior to the test period.

Finally, the popularity of *Pokémon GO* was considered an important characteristic because it influences social immersion. As a result of the game's popularity, users were more likely to encounter multiplayer scenarios (perhaps with friends) throughout the test period. It was believed multiplayer interactions would influence user immersion and engagement. Furthermore, a high number of participants would be likely to take part in the experiment, resulting in a larger participant distribution.

4.3.2 Just Dance Now

A typical session of *Just Dance Now* (Ubisoft, 2014) involves users dancing to a song of their choice by following on-screen directions. For each dance move, a graphic is displayed indicating how the user should move in order to score points (see Image 3.3A). In multiplayer scenarios, players can dance together using the same screen. *Just Dance Now* tracks movement via accelerometer and gyroscope sensor data from the phone that is held in the player's right hand.

Just Dance Now was chosen as the second exergame to test in this project. The most significant rationale for choosing this game is that it provides a unique form of exercise: dancing. This form of physical activity provides an extreme contrast to the walking in *Pokémon GO*.

The contrast between game elements and physical activity in *Just Dance Now* was noted to be significantly greater than that of *Pokémon GO*. Considering *Just Dance Now* is more focused on physical activity, the relationship between user enjoyment of the game itself and the physical activity provided by the game was believed to be more apparent than in *Pokémon GO*. This was important for selecting the two exergames in order to assess if players are enjoying the game or the physical activity provided by the game.

Another reason for choosing *Just Dance Now* was that the game is easy to test and play. The game is simple to set up with a feature that allows users to quickly join a room and connect to a monitor on which to play. The monitor shows images of the dance moves that the user needs to follow. The user will be able to understand what movements they must perform, regardless of their past dancing experience.

Lastly, *Just Dance Now* has essentially no social factor compared to *Pokémon GO*. Multiplayer scenarios are achieved in *Just Dance Now* when other players join the same room. However other players are not always physically present, unlike multiplayer encounters in *Pokémon GO*. This contrast was important in order to understand the impact of the social factor regarding exergames.

4.4 Game Testing Procedures

This section goes into detail about the procedures used in order to test the chosen exergames: *Pokémon GO* (Niantic Labs, 2016) and *Just Dance Now* (Ubisoft, 2014). Both of these procedures followed certain rules. The overall testing procedure was not to take over an hour to complete. Each testing session would begin with 30 minutes of gameplay that follows the instruction of certain rules, which are detailed below. Afterwards, the exergame enjoyment questionnaire was administered. Each user was provided with materials or equipment needed to play the games. In this case, that would be an Android Smartphone and, for *Just Dance Now*, another larger screen.

4.4.1 Pokémon GO

When testing *Pokémon GO*, there were not many rules restricting the players to specific walking paths, because the game is meant to be rather freeform. However, there are some rules that participants were still expected to follow. To not vary each player's experience too much, we limited the area in which the players could move to about a half mile radius. Also, while in this area, the player would need to visit ten "Pokéstops" in order to make sure that they were moving, as the game does not actually require the player to move.

4.4.2 Just Dance Now

For *Just Dance Now*, two users were placed in a room with a large screen that displayed the game's dance motions. They were both given an Android Smartphone and asked to play the game together for a duration of thirty minutes. A proctor would ensure that the users would be reasonably active throughout the testing period. The users were expected to play at least three unique songs.

Chapter 5: Questionnaire Development

A major goal of this project was to develop the Exergame Enjoyment Questionnaire in order to assess user enjoyment of exergames. This chapter elaborates on the development of the Exergame Enjoyment Questionnaire (Section 5.4). This chapter primarily focuses on providing rationales for questions that appear on the questionnaire. Furthermore, the format and design of the questionnaire is addressed in Section 5.4. Figure 5.4B shows the exact Exergame Enjoyment Questionnaire administered to users throughout this project.

5.1 Questions from Existing Questionnaires

This section describes rationales for questions that were influenced by existing GEQ and IEQ questionnaires. The questions that appear in this section were specifically targeted towards assessing user engagement and immersion. Throughout the development of the questionnaire, two existing questionnaire studies were used to derive questions for use in the Exergame Enjoyment Questionnaire. The two studies involved were, "The development of the Game Engagement Questionnaire: A measure of engagement in video game-playing" (Brockmeyer et al., 2009) and "Measuring and Defining the Experience of Immersion in Games" (Cox et al., 2008).

GEQ Question #1 (Brockmeyer et al., 2009)	**Exergame Enjoyment Questionnaire**
I lose track of time.	I felt like I lost track of time while playing.

This question focuses on user engagement via the component of absorption (Section 2.1.1). By losing track of time, a user stops paying attention to their surroundings. As a result, it could be considered that the user is completely absorbed and focused on the game.

GEQ Question #19 (Brockmeyer et al., 2009)	**Exergame Enjoyment Questionnaire**
I feel like I just can't stop playing.	I did not feel like I wanted to keep playing.

This question was selected because it allowed the questionnaire to acquire an alternative perspective on enjoyment. It was believed that this question, in combination with the rest of the questionnaire, would provide a holistic perspective on enjoyment. For example, if a user disagreed to this question and their answers to other questions showed high engagement and immersion, then it would be concluded that the user expressed a high level of enjoyment overall.

IEQ Question #2 Appendix A (Cox et al., 2008)	**Exergame Enjoyment Questionnaire**
I did not feel any emotional attachment to the game.	I felt emotionally attached to the game.

This question was chosen in order to assess extreme levels of immersion through emotional involvement. It was believed that many users would not develop an emotional attachment to a game unless they experienced an extremely high level of immersion. This question was important to this project considering users had the potential to become emotionally attached to music that was part of the game, such as in *Just Dance Now!*.

IEQ Question #22 Appendix A (Cox, 2008)	**Exergame Enjoyment Questionnaire**
I was aware of surroundings.	I felt a strong sense of being in the world of the game to the point that I was unaware of my surroundings.

This question also focused on assessing extreme levels of immersion. Players that lost a sense of reality while playing the game, such that they were no longer paying attention to events that occurred around them, were believed to be fully immersed in the game. This question furthered the ability of the questionnaire to assess extremely high levels of immersion.

IEQ Question #2 Appendix B (Cox et al., 2008)	**Exergame Enjoyment Questionnaire**
To what extent did you feel you were focused on the game?	I was focused on the game.

This question was used to gain insight on how invested the user was in the game. If a user agreed to this question, a high level of immersion was more likely to have occurred for that user. In contrast, users that disagreed or answered neutral to this question were considered to have experienced low, if any, immersion in the game.

IEQ Question #24 Appendix B (Cox et al., 2008)	**Exergame Enjoyment Questionnaire**
To what extent were you interested in seeing how the game's events would progress?	I did not feel a desire to make progress in the game.

This question was used in the questionnaire in order to determine if the user was actively trying to make progress in the game. If a user disagreed with this statement then that user would be considered immersed in the game. This comes as a result of the user consciously attempting to reach a goal or achievement within the game.

5.2 Questions Inspired by the PACES Questionnaire

In addition to asking questions about the game element of exergame enjoyment, it is important to also analyze the exercise component of exergame enjoyment. The *PACES* questionnaire was used as a source of physical activity-related questions. Table 5.2 gives examples of statements from the *PACES* questionnaire that have been altered in order to apply them to exergames, specifically.

Table 5.2: PACES Statements Modified for an Exergame Questionnaire

Original Statement from the *PACES* Questionnaire When I am physically active...	**Exergame Questionnaire**	**Rationale**
It feels good.	The exercise in this game made me feel good.	Exercise provided in the games should give users a sense of physical well-being.
It is very exciting.	I felt excited about the physical activities in the game.	If the user is not excited by the game or if they find the game boring, then the exergame does not accomplish its purpose of getting inactive people to exercise more.
I get something out of it.	I felt that playing the game was beneficial for my physical well-being.	The user should believe that their physical activity in the game is of benefit to their physical well-being.
I feel as though I would rather be doing something else.	I would rather not be exercising, even though the exercise was accompanied by game elements.	If the user decides that they would not play the game voluntarily, outside of the testing period, then that means the game is not enjoyable enough.

5.3 Original Questions

This section describes rationales for questions that are original, not taken from other questionnaires. These questions attempt to test for certain thought processes and behaviors not tested in the previously created questions. Questions were created for the categories of Intrinsically Rewarding Activity, Control and Exercise, we did not create any new questions for Immersion. Refer to Sections 2.2.2 and 5.4 for category descriptions.

I felt that this game provided an enjoyable challenge.

This question was created to test intrinsically rewarding activity. An intrinsic reward is defined as, "an outcome that gives an individual person satisfaction such as that derived from a job well done" (businessdictionary.com). It tests if a user feels that the actual gameplay of the game is enjoyable or challenging. This question is important since it the most direct question at asking if a player if they found the game rewarding

I felt a sense of accomplishment from playing the game.

This question was also created in order to test intrinsically rewarding activity. Unlike the previous question it does not directly ask whether or not the user enjoyed the game, but rather what they accomplished. It is important to measure enjoyment without using the word "enjoyment", in the cases where people have preconceptions on what is enjoyment.

I felt that It was easy to familiarize myself with the game controls.

Control is defined as, “the power to influence or direct people’s behavior or the course of events” (oxforddictionaries.com). The goal of this question was to measure how difficult the user thought it was to control the game. “Controls” refers to user interface interactions and the user experience within the game. Examples targets in this question were buttons, sliders, selectors, and other gameplay options available in the game’s user interface.

I felt that it was difficult to understand how the game works.

This question was used to measure control of the game. Instead of using the word “control” directly, which has heavy implications on the actual buttons and interface you interact with, “how the game works” implies interacting with the game as a whole. Additionally, it is phrased negatively. This is an important question, as sometimes a game’s physical controls may be lacking, however the game interaction in general is creative and intuitive.

I felt in control of the game.

This question was created to measure control of the game in the most basic way. This question tests the overall control. A very basic question to compare the others to. Much like the question above, it does not directly refer to the physical controls of a game.

I felt that the game reacted quickly to my actions.

This question was created to measure control of the game. This question gathers whether or not the player was able to do what they wanted in the game. When they made an action, was that action reflected quickly and accurately in the game. This is often an important aspect that leads to the feeling of control within a game.

I consider playing the game “exercise”.

This question was created to gather information of the exercise factor within an exergame. An important aspect of exergames is that they require the user to use their body in some way as part of the game. This question tests if the specific user playing the game considers the movements and actions they do to be exercise. The goal of this question is to evaluate the movement in the game.

I felt that the game would have been more enjoyable without physical activity.

This question was created to gather information of the exercise factor within an exergame. If a player enjoyed or did not enjoy the game, it is important to find what aspects they enjoyed or lead to such enjoyment. In an exergame the most basic of two categories is the physical activity and the game elements. The goal of this question was to see if the enjoyment the game bore was from the exercise element or the game element.

I would prefer that this physical activity was not accompanied by game elements.

This question was created to gather information of the exercise factor within an exergame. It is similar to the last question but instead of inquiring about the physical activity, it target’s the game elements. This is to see whether or not the physical activity or the game elements appealed more to the user.

I felt that the physical activity was too intense for me.

This question was utilized to measure the level of exercise provided by the exergame through the user’s perspective. By assessing the intensity of physical activity provided by the exergame, it was possible to determine if users enjoyed the physical activity.

5.4 The Exergame Enjoyment Questionnaire

This section describes the categorization of the final questionnaire(table 5.4A) and how the questionnaire is presented to the user.

Table 5.4A: Exergame Enjoyment Questionnaire Categories

Immersion	1. I did not feel like I wanted to keep playing. 2. I felt like I lost track of time while playing. 3. I felt a strong sense of being in the world of the game to the point that I was unaware of my surroundings. 4. I felt emotionally attached to the game. 5. I was focused on the game.
Intrinsically Rewarding Activity	1. I felt that this game provided an enjoyable challenge. 2. I felt a sense of accomplishment from playing the game. 3. I did not feel a desire to make progress in the game.
Control	1. I felt that It was easy to familiarize myself with the game controls. 2. I felt that it was difficult to understand how the game works. 3. I felt in control of the game. 4. I felt that the game reacted quickly to my actions.
Exercise	1. I consider playing the game “exercise”. 2. I felt that the game would have been more enjoyable without physical activity. 3. The exercise in this game made me feel good. 4. I felt excited about the physical activities in the game. 5. I felt that playing the game was beneficial for my physical well-being. 6. I would rather not be exercising, even though the exercise was accompanied by game elements. 7. I would prefer that this physical activity was not accompanied by game elements. 8. I felt that the physical activity was too intense for me.

5.4.1 Categories

The questionnaire was created to focus on four main categories: “Immersion”, “Intrinsically Rewarding Activity”, “Control”, and “Exercise”. The first three sections contain questions that access player enjoyment of game elements in the exergame. These three categories are the three major factors in measuring game enjoyment (Chapter 2). The section, “Exercise” contains questions that assess player enjoyment of the physical activity provided by the game. Considering player enjoyment of exercise is the main focus of this project, this section contains the most questions, making up eight out of the twenty questions on the questionnaire. The other three sections contains twelve questions altogether.

5.4.2 Format

Although the questionnaire is divided into four categories, when presented to users, the order of the questions is randomized to avoid consecutive questions of the same category. This prevents participants from selecting similar answers for questions that are in the same section with minimal cognitive effort or without thinking about their responces.

Users answer the questions using a five-point Likert scale: strongly disagree, disagree, neutral, agree, strongly agree. (Chapter 3.5) Questions 4, 6, 10, 11, 13, 19, 20 are negatively phrased. User’s answers to these questions are then flipped, in order to maintain consistently positive scoring of the questionnaire (i.e. higher total scores indicate more enjoyment of the exergame). The final questionnaire presented to users is shown in figure 5.4B.

Figure 5.4B: The Exergame Enjoyment Questionnaire As Administered

Run for Fun: Exergame Enjoyment Questionnaire

For each of the statements below, please circle how much you agree or disagree with the statement.

1. I felt excited about the physical activities in the game.

Strongly Disagree *Disagree* *Neutral* *Agree* *Strongly Agree*

2. The exercise in this game made me feel good.

Strongly Disagree *Disagree* *Neutral* *Agree* *Strongly Agree*

3. I felt like I lost track of time while playing

Strongly Disagree *Disagree* *Neutral* *Agree* *Strongly Agree*

4. I felt that it was difficult to understand how the game works.

Strongly Disagree *Disagree* *Neutral* *Agree* *Strongly Agree*

5. I was focused on the game.

Strongly Disagree *Disagree* *Neutral* *Agree* *Strongly Agree*

6. I felt that the game would have been more enjoyable without physical activity.

Strongly Disagree *Disagree* *Neutral* *Agree* *Strongly Agree*

7. I felt that it was easy to familiarize myself with the game controls.

Strongly Disagree *Disagree* *Neutral* *Agree* *Strongly Agree*

8. I felt emotionally attached to the game.

Strongly Disagree *Disagree* *Neutral* *Agree* *Strongly Agree*

9. I consider playing the game “exercise”.

Strongly Disagree *Disagree* *Neutral* *Agree* *Strongly Agree*

10. I felt that the physical activity was too intense for me.

Strongly Disagree *Disagree* *Neutral* *Agree* *Strongly Agree*

11. I did not feel a desire to make progress in the game.

Strongly Disagree *Disagree* *Neutral* *Agree* *Strongly Agree*

12. I felt a strong sense of being in the world of the game to the point that I was unaware of my surroundings.

Strongly Disagree *Disagree* *Neutral* *Agree* *Strongly Agree*

13. I would rather not be exercising, even though the exercise was accompanied by game elements.

Strongly Disagree *Disagree* *Neutral* *Agree* *Strongly Agree*

14. I felt that playing the game was beneficial for my physical well-being.

Strongly Disagree *Disagree* *Neutral* *Agree* *Strongly Agree*

15. I felt that this game provided an enjoyable challenge.

Strongly Disagree *Disagree* *Neutral* *Agree* *Strongly Agree*

16. I felt a sense of accomplishment from playing the game.

Strongly Disagree *Disagree* *Neutral* *Agree* *Strongly Agree*

17. I felt that the game reacted quickly to my actions.

Strongly Disagree *Disagree* *Neutral* *Agree* *Strongly Agree*

18. I did not feel like I wanted to keep playing.

Strongly Disagree *Disagree* *Neutral* *Agree* *Strongly Agree*

19. I would prefer that this physical activity was not accompanied by game elements.

Strongly Disagree *Disagree* *Neutral* *Agree* *Strongly Agree*

20. I felt in control of the game.

Strongly Disagree *Disagree* *Neutral* *Agree* *Strongly Agree*

5.5 Additional Questionnaires

In addition to the Exergame Enjoyment Questionnaire, test participants will also be given a Pre-Exergame Questionnaire. This additional questionnaire is intended to provide context for the results of the Exergame Enjoyment Questionnaire so that underlying relationship between certain sub-populations and certain answers could easily be discovered. If there is a correlation between exergame enjoyment and some other factor not captured by the Exergame Enjoyment Questionnaire, such as gender, age, or physical activity experience, then these questionnaire components will help display that relationship. This additional questionnaire will also reveal whether if the Exergame Enjoyment Questionnaire is invalid for a particular demographic. Table 5.5 below describes the questions in the Pre-Exergame Questionnaire and the rationale behind them.

Table 5.5 Pre-Exergame Questionnaire

Question	**Rationale**
Participant Gender: _____	There could be a potential difference in exergame enjoyment for men and women.
Participant Age: _____	Most or all of the test participants were expected to be college students of ages 18-22. Even though there is little variance in this predicted age distribution, there could still be a noticeable change in exergame enjoyment as people grow throughout college. Future studies may also target non-college populations.
Do you know what is an exergame?	If a participant is previously familiar with exergames, then they may be more likely to understand the value or purpose of exergames. There may be a connection between a participant's prior knowledge of exergames and their score on the Exergame Enjoyment Questionnaire.
Have you heard of or played either of the games: *Just Dance Now* or *Pokémon GO*?	If a player's enjoyment of a game is influenced by whether or not they have already played it previously, then this is a factor that should be considered when observing their Exergame Enjoyment Questionnaire results.
What, if any, mobile games have you played?	There may be a connection between a player's taste in mobile games and their taste in exergames. The answers to this question could explain how the participant feels about the game elements of *Pokémon GO* and *Just Dance Now*.
How many hours a week would you say you engage in physical activity?	There could be a difference in exergame enjoyment for people that exercise regularly and for people who live sedentary lifestyles. For instance, sedentary people may dislike all physical activity, even if it is gamified. In particular, if exergames can appeal to those who are inactive, then it can help motivate them to be healthier.
How many hours a week would you say that you play exergames?	There could be a distinction in exergame enjoyment for people that play exergames regularly and those who do not. In particular, it would be interesting to determine if regular exergame players enjoy *Pokémon GO* and *Just Dance Now*.

Chapter 6: Focus Groups

This chapter describes the final version of the testing procedures for both games. Additionally, this chapter discusses the overall procedure and methodology revisions that resulted from conducting focus groups.

6.1 Procedure

The general testing procedure for participants is described in the list below with specific gameplay scenarios described in Sections 6.1.1 and 6.1.2.

1. Sign the informed consent form
2. Administer pre-questionnaire
3. Play the game (Pokémon GO or Just Dance Now) for thirty minutes
4. Administer the Exergame Enjoyment Questionnaire
5. Discuss the focus group conversation questions (Section 6.1.3)

6.1.1 Pokémon GO

Participants received phones with *Pokémon GO* already open with a brand new Pokemon Go account. Participants were instructed to go through the tutorial and then went outside to catch Pokémon and visit Pokéstops for thirty minutes. Additionally, participants were given an instruction sheet with a campus map, passcode for the phone, and a list of objectives that they should attempt to complete within thirty minutes. The campus map, shown in Figure 6.2.1, shows the boundary that participants should remain within throughout the test. This forced the participants to remain in the WiFi range of the WPI campus in order to prevent disconnection issues. The phone passcodes were provided in case the participants turned the phone off accidentally. The list of objectives encourage participants to walk more and to have a full experience of the game.

6.1.2 Just Dance Now!

Participants played *Just Dance Now!* for thirty minutes after receiving a phone that had the game opened and connected to a "Dance Room". Participants would play the game in a wide classroom with enough space for two or three people to dance simultaneously. Additionally, a projector displayed the "Dance Room" that the participants would see the dance moves required for each song. Furthermore, this project utilized the VIP version of *Just Dance Now!* Such that players could access all of the songs in the game. Lastly, participants were instructed to play the game at their own pace and intensity, but were expected to play the game without breaks for the entire thirty minutes.

6.1.3 Focus Group Questions

In order to determine if there were any outstanding problems with the Exergame Enjoyment Questionnaire, discussions were held with participants after playing each game. This process was utilized in order to gather the participants' opinions on the project's procedure and the Exergame Enjoyment Questionnaire. Based on the participants responses, it was possible to determine where problems exist in the methodology and the Exergame Enjoyment Questionnaire itself.The questions listed below were used as simple prompts, in order to initiate discussion between the participants and the researchers.

1. How do you feel about the game?
2. Would you play this game again?
3. What do you think about the physical activity in this game?
4. While playing, did you reach any goals, achievements or accomplishments?
5. Was there anything you did not like about the game?
6. Did anything interrupt your gameplay?
7. Did you have any problems while taking the questionnaire?

6.2 Methodology Revisions

This section justifies aspects of the focus group testing procedure. The first few focus groups demonstrated problems with the initial methodology that required revision. Four procedural changes - two for each game - are described below. These adjustments have been made for the purpose of both giving players the *same* experience and also a more *enjoyable* experience.

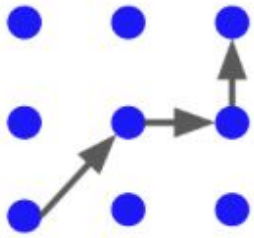
6.2.1 Pokémon GO

Dealing with WiFi Disconnections: It was apparent from the first few Pokémon GO focus groups that the mobile phones all had poor connections to campus WiFi. During just a thirty minute long testing period, a participant could expect to disconnect at least four or five times. Furthermore, the WiFi would disconnect entirely if the player drifted too far from the center of campus. In order to prevent this from ruining the gameplay experience, an instructions sheet was given to every person (shown in Figure 6.2.1). This sheet of paper includes a map of campus with a red boundary to display where the WiFi connection will significantly weaken. It also includes objectives for the user to try completing as well as the password for the phone in case they accidentally turn it off while they are playing.

Normalizing Leveling Up Events: From the first few Pokémon GO focus groups, it also became clear that players enjoyed 'leveling up' during the game. The problem was that the rate of leveling up slows down as the level increases; for example, the game requires more play to move up from level 8 to level 9 than it does to move up from level 2 to level 3. In addition, higher levels provide the user with more opportunities; for example a user can participate in gym battles only if they reach level 5. For these two reasons, it became evident that different users would have different experiences unless they all began from the same starting point. To maintain test consistency and remove the 'leveling up' variable, the methodology was adjusted so that every user would play the game on a brand new account rather than from where the last person left off.

Figure 6.2.1 Pokémon GO Focus Group Instructions Sheet

Phone Password



Instructions

- Return within a half hour of the starting time.
- Stay within the red boundary shown above.
- Visit at least 10 PokéStops.
- Catch Pokémons that appear.

6.2.2 Just Dance Now

During the first few testing sessions of Just Dance Now, we limited the users to just two songs. This was for two reasons; the game had two songs by default (additional songs would require purchase or extensive gameplay), and forcing all the players to play the same songs enabled consistency between tests. However something quickly became apparent; users did not enjoy repeatedly playing the same two songs for an entire 30 minute period of time. Many players commented that they would have enjoyed the experience more if they had access to more songs. In response to this complaint, future test participants were given access to all the songs in the game and could choose whichever song they wished to play next, be it a repeat or a brand new song.

Another factor that changed was the number of players tested simultaneously. At first, we had participants play the game alone. Aside from the test proctors, there were no other individuals in the room with the user. In these cases the user sometimes felt awkward or would not put much effort into playing. We hypothesized that this was because Just Dance Now is a social game and it was therefore strange to play without others. During the future test sessions, we changed the testing methodology so that the game would only be played with two or three users simultaneously.

6.3 Questionnaire Revisions

This section details changes made to the Exergame Enjoyment Questionnaire. In total, four questions were rephrased and one question was removed entirely as a result of the focus group process. Versions of the questions before and after the focus groups are shown in tables below. Words or phrases that were considered ambiguous or vague are bolded in each question.

Question #13

Before Focus Groups	**After Focus Groups**
I would rather not be exercising, even though the exercise was **'gamified'**	I would rather not be exercising, even though the exercise was accompanied by game elements.

This question was modified in order to clarify the intent of the question. The word, "gamified" was considered confusing to the players, which affected their ability to properly interpret the meaning of the question. When asked about the question, a player referred to 'gamified' as, "Yeah, I'd never heard that before." The word 'gamified' was changed to the phrase, "accompanied by game elements" to more clearly define that the question was asking about the exercise that was part of the game.

Question #14

Before Focus Groups	**After Focus Groups**
I felt that I **benefitted** from playing the game.	I felt that playing the game was beneficial for my physical well-being.

The word, "benefitted" made this question confusing to multiple players, considering that the meaning of the word was overall subjective and somewhat ambiguous. During a focus group discussion, a player said, "I just didn't really know what it means. I was just thinking maybe it meant that I benefitted from the exercise. But I wasn't really sure." Overall, "benefitted" was considered an issue because players could not immediately determine the meaning of the question. For that reason, the word "benefitted" was replaced with the phrase, "beneficial for my physical well-being." This wording was considered to provide enough detail such that players could determine the meaning of the question.

Question #15

Before Focus Groups	**After Focus Groups**
I cared about winning in the game.	[Removed]

Throughout the focus group process, this question was discovered to provide unnecessary information. Additionally, it was deemed confusing because the word, "winning" was very subjective. Players were unable to determine how one could "win" a game such as Pokémon GO, when there exists no such objective in order to "win" the game. A player referred to this by adding, "There's no clear cut goal for winning. You make your own goal, I think, for this game." Overall, this question was considered confusing and provided information that was essentially meaningless regarding user engagement and immersion, hence its removal from the questionnaire.

Question #17

Before Focus Groups	**After Focus Groups**
I felt that the game was **responsive**.	I felt that the game reacted quickly to my actions.

The meaning of this question was determined to be subjective because of the word, "responsive". A player described their interpretation of this question as, "I felt that the game was responsive, though - I didn't understand that." It was decided that the term, "responsive" should be changed to a more clear phrase, "reacted quickly to my actions." This made the meaning of the question easier to interpret.

Question #19

Before Focus Groups	**After Focus Groups**
I would prefer that this physical activity was not '**gamified**'.	I would prefer that this physical activity was not accompanied by game elements.

This question was modified for the same reasons as Question #13. The issue with this question was again, the word "gamified". Refer to Question #13 in this section for more details.

Chapter 7: Questionnaire Validation

The objective for validating the EEQ was to provide evidence that the questionnaire generates reliable user exergame enjoyment data. This chapter outlines the Exergame Enjoyment Questionnaire validation process and results. Furthermore, the coding process and methodology by which the EEQ was validated is explained. Lastly, this chapter discusses the effectiveness and accuracy by which the EEQ assessed user enjoyment of exergames.

7.1 Coding Procedure

Coding Categories

In order to validate the questionnaire, providing evidence that it accurately assessed user exergame enjoyment, it was necessary to "code" the user's responses from the focus group discussion. Coding is the process of assigning an identifier to a value or response in order to classify the types of responses. The coding process in this project involved picking out words or phrases in the user answers to questions on the EEQ which were similar to their responses in the focus group discussion. For each of the twenty questions on the Exergame Enjoyment Questionnaire, one of three categories was applied to indicate Agreement, Disagreement, or No Evidence. These categories are described below.

1. **Agreement**

This category indicated that a user's response in the focus group discussion was consistent with their response to a given question on the Exergame Enjoyment Questionnaire.

Example

- Exergame Enjoyment Questionnaire Question: I consider playing the game "exercise".
- User's Answer to Question: Strongly Agree
- Focus Group Discussion: "By halfway through I was actually starting to feel a workout to it"

2. **Disagreement**

This category signified that a user response in the focus group discussion was inconsistent with or contradicted their response to a given question on the Exergame Enjoyment Questionnaire.

Example

- Exergame Enjoyment Questionnaire Question: I felt that the game reacted quickly to my actions.
- User's Answer to Question: Agree
- Focus Group Discussion: "[only thing I didn't like] it was delayed because of the wireless connection"

3. **No Evidence**

This category signified that there was insufficient evidence obtained from the focus group discussion that could be used to indicate Agreement or Disagreement.

Example

- Exergame Enjoyment Questionnaire Question: I consider playing the game "exercise".
- User's Answer to Question: Strongly Agree
- Focus Group Discussion: The topic was not discussed or the participant did not provide any evidence in order to make a decision on Agreement or Disagreement.

Coding Guidelines

Throughout the validation process, each of the four researchers involved in the project separately coded a number of participant responses. Therefore, it was necessary to develop a standard set of guidelines in order to maintain consistency between codings. These guidelines were used to interpret user EEQ responses with what the user said in the focus group discussion. For example, if a user answered, "Neutral" to any question on the EEQ, the researcher would refer to the corresponding guideline in order to correctly code the user's response relative to the user's focus group discussion responses. Several guidelines were developed, each of which is described below. {Strongly Disagree, Disagree, Agree, and Strongly Agree} user responses followed the same guideline, as it was considered extremely difficult to gather evidence to distinguish between answers that were "Strong". A {Neutral} response required its own guideline considering there were multiple cases to consider in order to validate this response.

Lastly, a guideline was created to handle question "consistency". This guideline was used to validate questions that should have had similar user responses on the EEQ or are in the same category (Section 5.4.1). Some questions should have similar user responses considering there are questions that are essentially the same, but phrased differently. Furthermore, questions that are in the same category should have had similar user responses depending on what they were targeting. Further details on this guideline are outlined below.

1. Neutral Responses

If a user responded to a question on the EEQ with a neutral response, then the question was marked as "Agreement" or "Disagreement" according to the following rules:

1. *Non-Definitive or Double Sided*
    a. If the user indicates a non-definitive opinion ("I thought it was okay") or opinions that are contradictory ("I enjoyed the exercise but it was boring"), then mark the question as "Agreement". This case occurs when the user either truly does not hold an opinion on the topic, or gives contradicting evidence which would indicate that the user is split between "Agree" and "Disagree".
2. *One-Sided*
    a. If the user indicates a definitive opinion about a question in the focus group discussion, then mark the question as "Disagreement". A definitive opinion does not match with a neutral answer.

2. Non-Neutral Responses (Strongly Disagree, Disagree, Agree, Strongly Agree)

If a user responded to a question on the EEQ with a non-neutral response (Strongly Disagree, Disagree, Agree, Strongly Agree) and the user gave any indication of a definitive opinion about the question in the focus group discussion, then the question was marked as an "Agreement" or "Disagreement" appropriately.

3. **Consistency Between Questions**

If a question had a user response that contradicted that of another question it should be consistent with, then the question was marked as "Disagreement". Otherwise, if the questions were consistent, the question was marked as "Agreement". The following questions were considered to be consistent with each other:

- Questions 6, 13, 19
- Questions 2, 9, 14
- Questions 4, 7

Peer Reviews

In order to ensure that researchers coded participant responses similarly to each other, a peer review process was utilized to ensure consistency between codings from different researchers. First, four different participant responses were coded by each researcher. A discussion was held between the researchers about the differences in their obtained coding values. This resulted in a final validation document for each of the four coded participants. Subsequently, the four researchers split into groups of two, where each group coded two more participants. A discussion was again held between the two coders of each participant's responses in order to create a final validation document. It was believed that at this point that the researchers were now able to individually code participant responses while maintaining consistency with the other researchers.

7.2 Validation Results

After conducting user tests, 38 user responses (Out of a total 38 participants) were coded with the process described in Section 7.1. Figure 7.2A displays the distribution of Agreements and Disagreements for each question on the EEQ. The number of Disagreements increase to the left of the chart, while the number of Agreements increase to the right of the chart. Each horizontal bar corresponds to the Disagreements and Agreements for the given question. This chart provides a holistic perspective on questions that did and did not accurately assess user enjoyment of exergames. Questions with a high percentage of Disagreements were considered problematic, while those with a high percentage of agreements were considered to be acceptable. Figure 7.2B provides the distribution of "No Evidence" for each question. Questions with a high number of "No Evidence" were considered to be extremes, such that they were difficult to indirectly ask users about in the focus group discussion. Furthermore, a high number of "No Evidence" was indicative of a problem with the focus group process itself, not necessarily the question. It could have been the case that the researchers developed too few questions to ask, or the questions themselves were flawed and did not provide useful material in order to code user responses.

In order to identify problems within specific categories of the EEQ, a separate distribution of Agreement and Disagreement was created for each category. Figures 7.2C to 7.2F illustrate the distributions for each category, including: Physical Activity, Immersion, Intrinsic Reward, and Control. These distributions make it easier to identify problematic questions. In each category's chart, it is possible to find which questions encountered issues with accurately assessing user enjoyment. For example, in Figures 7.2A and 7.2C, question nine has a much higher number of disagreements compared to the other questions in the same category. This indicates that question nine was most likely flawed.

Figure 7.2G provides a chart of "Agreement" Percentages for each question on the EEQ. The value for each question was calculated using the following equation:

$$Agreement\ Percentage = \frac{Agreements}{(Agreements + Disagreements)}$$

By calculating the percentage of agreements for each question, it became easier to determine the most flawed questions. Those with a high agreement percentage were considered valid and those with a low agreement percentage were considered flawed.

Finally, figures 7.2H and 7.2I provide the Agreement and Disagreement distribution specific to *Just Dance Now!* and *Pokémon GO*, respectively. These distributions emphasize the questions that could have been more valid for one game as opposed to the other. Additionally, they provide perspective as to which game contributed the most to the overall number of agreements and disagreements.

Refer to Section 9.1 for conclusions on the validity of the EEQ.

<u>Note: Figures 7.2A - 7.2G include data from both Just Dance Now! and Pokémon GO.</u>

Figure 7.2A: Number of Agreements and Disagreements for each question. Sorted by Disagreements, greatest to least (worst to best).

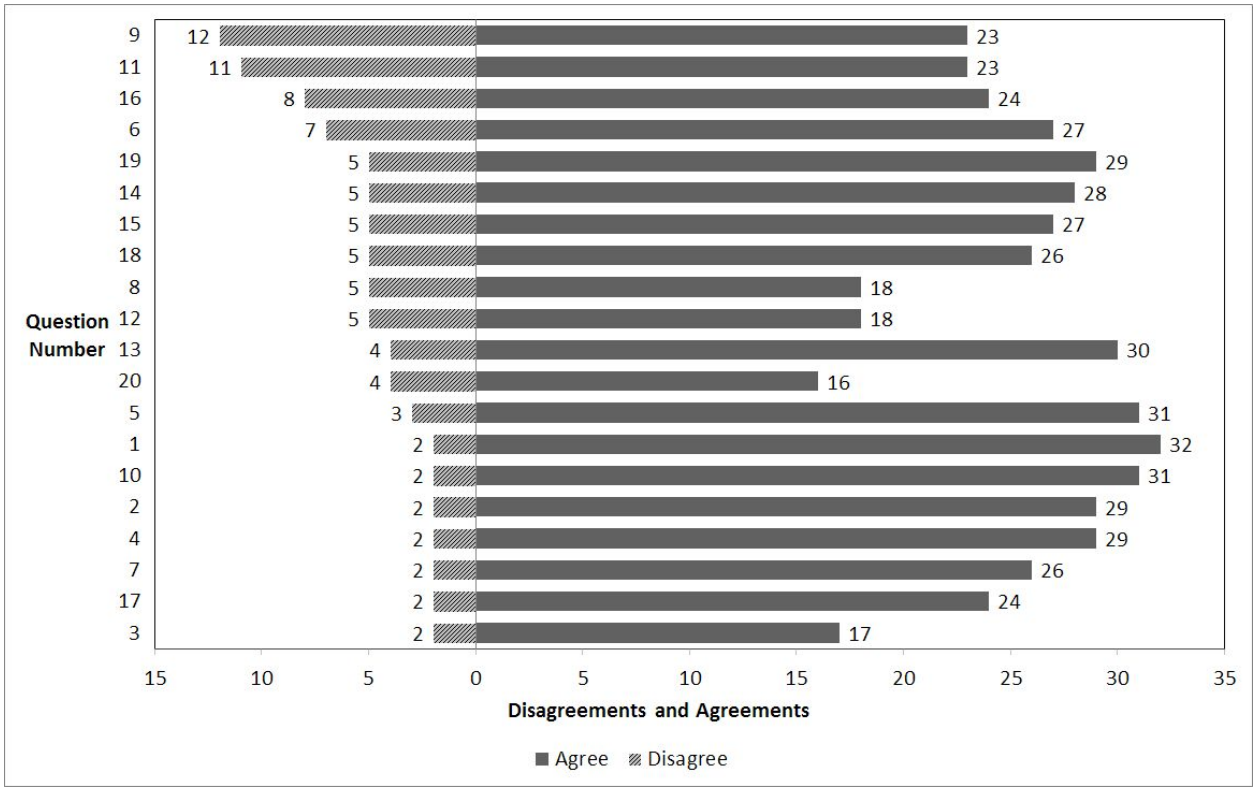


Figure 7.2B: Number of "No Evidence" occurrences for each question. Sorted by number of No Evidence greatest to least (worst to best).

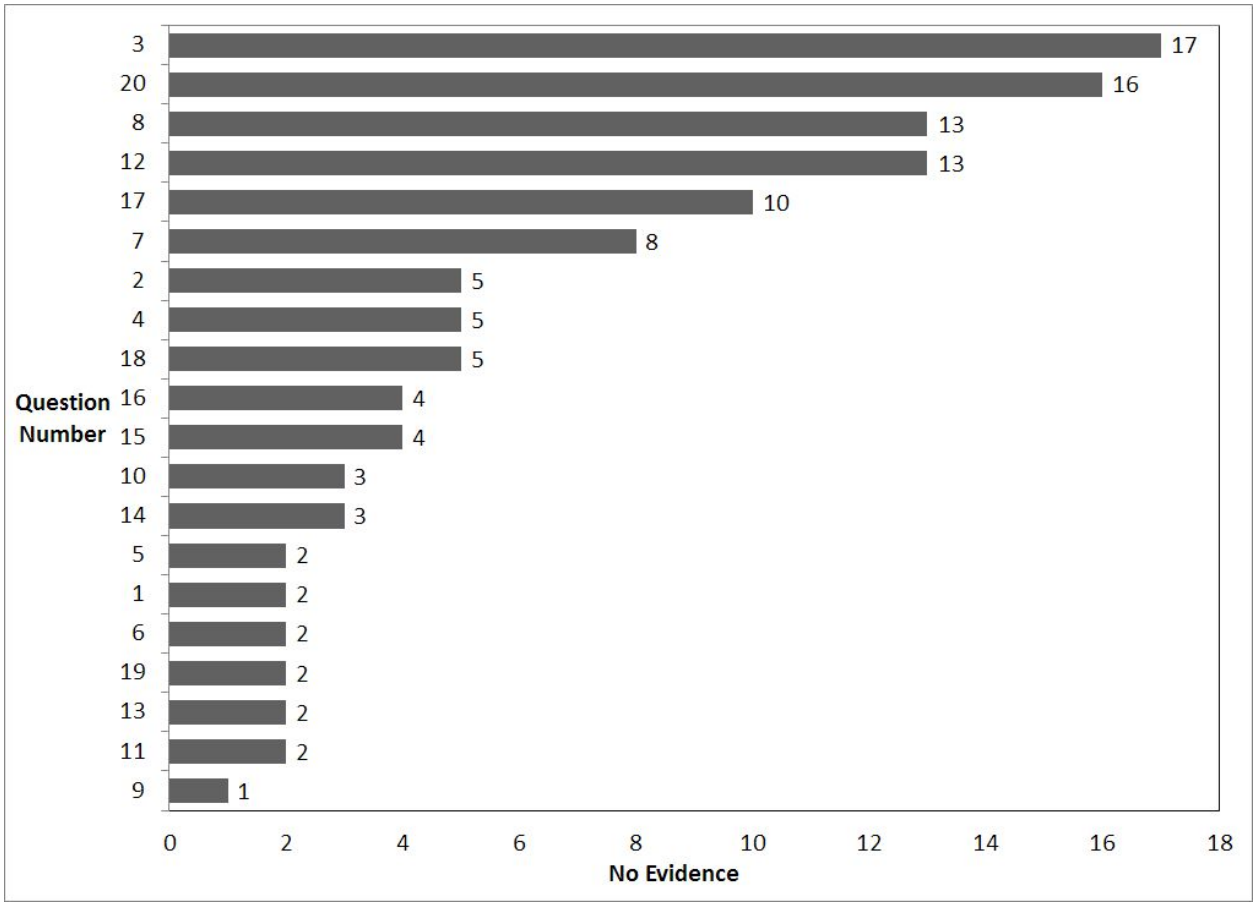


Figure 7.2C: Number of Agreements and Disagreements for each question in the ***Physical Activity*** *category. Sorted by Agreements, greatest to least (best to worst).*

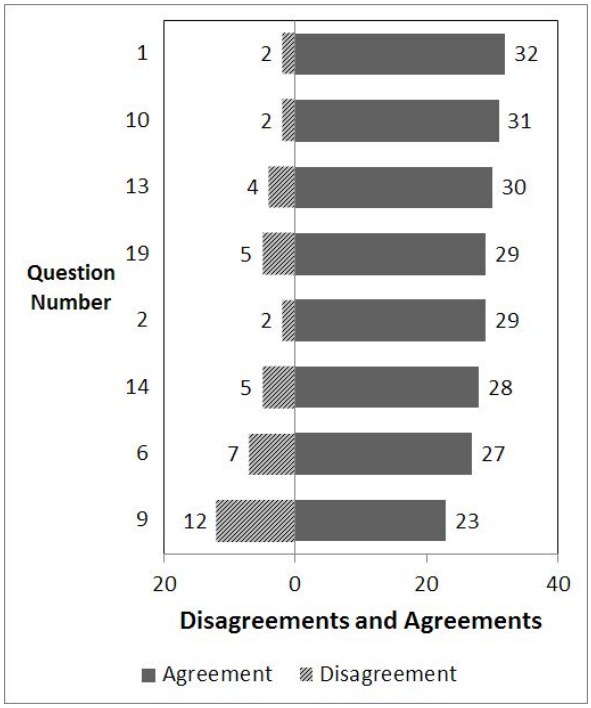


Figure 7.2D: Number of Agreements and Disagreements for each question in the ***Immersion*** *category. Sorted by Agreements, greatest to least (best to worst).*

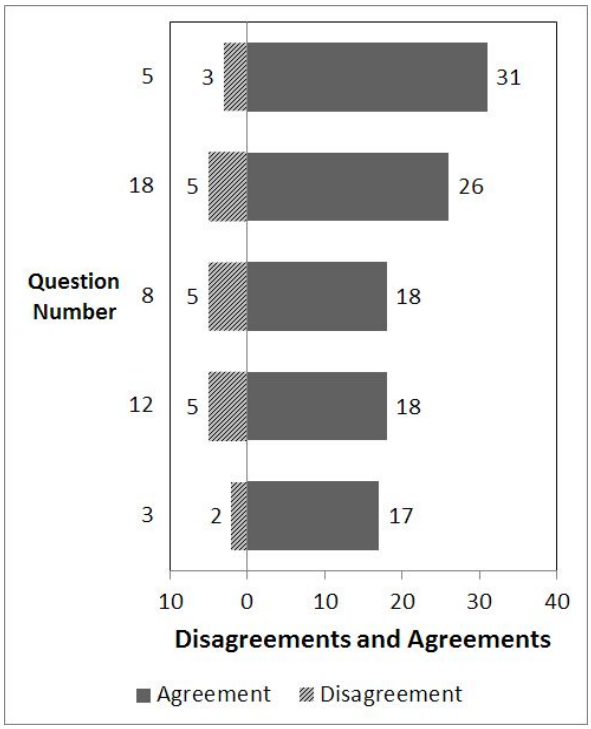


Figure 7.2E: Number of Agreements and Disagreements for each question in the ***Control*** *category. Sorted by Agreements, greatest to least (best to worst).*

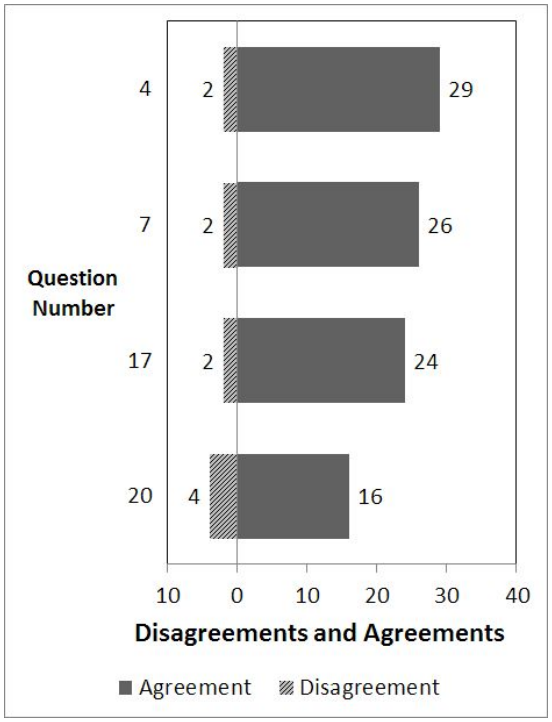


Figure 7.2F: Number of Agreements and Disagreements for each question in the ***Intrinsic Reward*** *category. Sorted by Agreements, greatest to least (best to worst).*

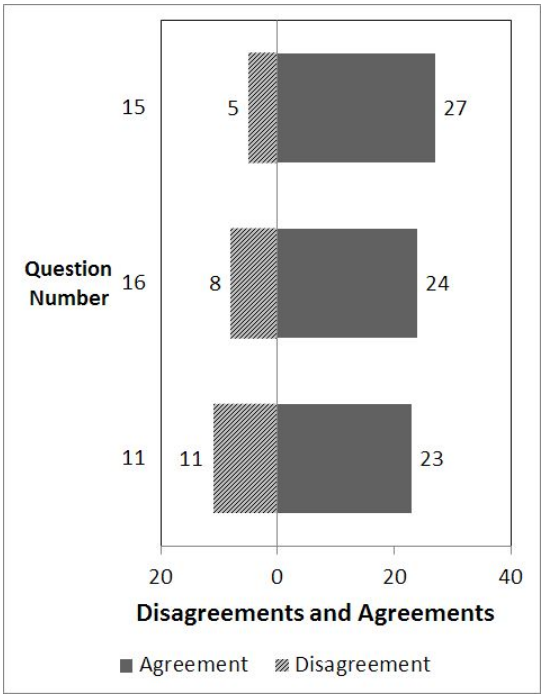
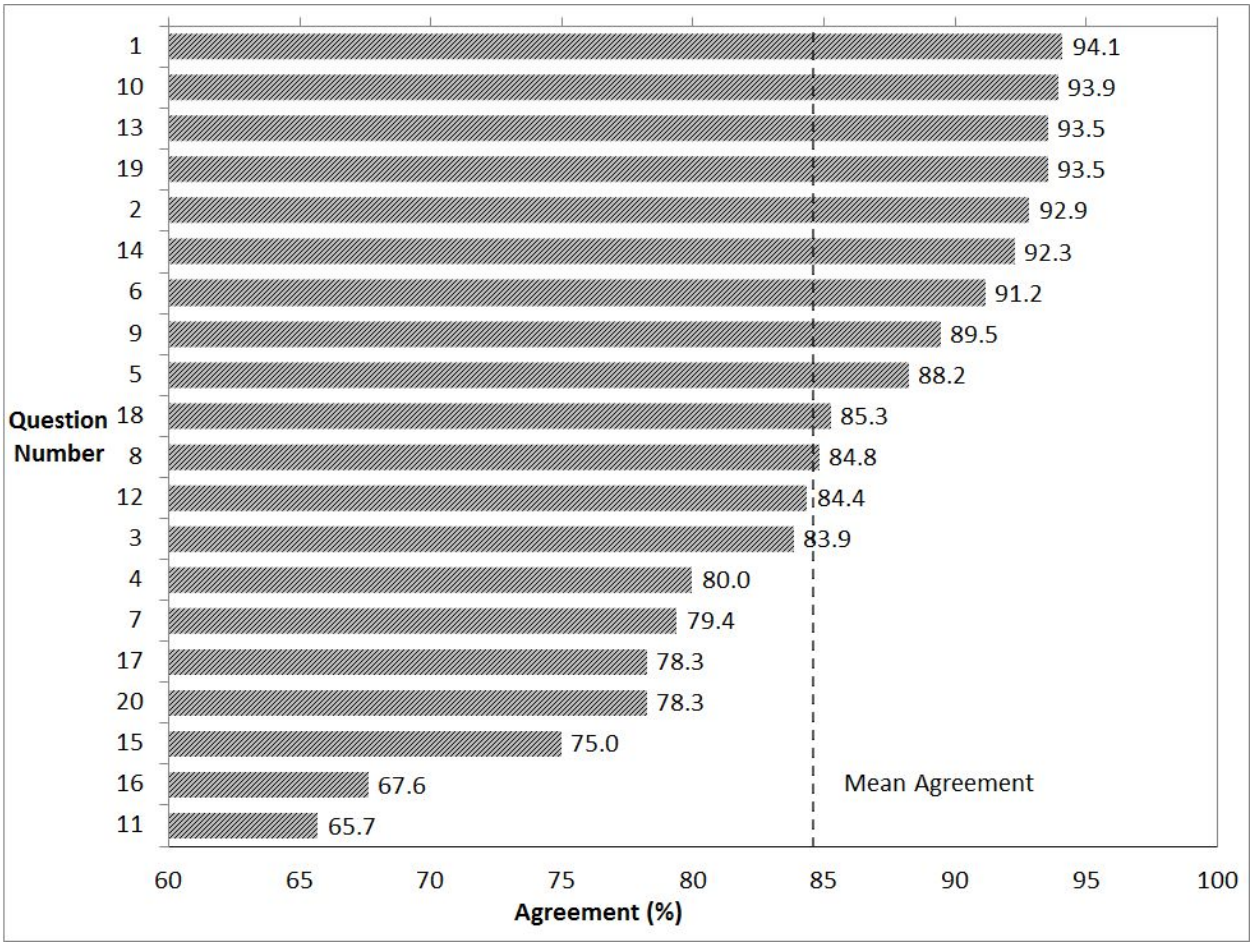
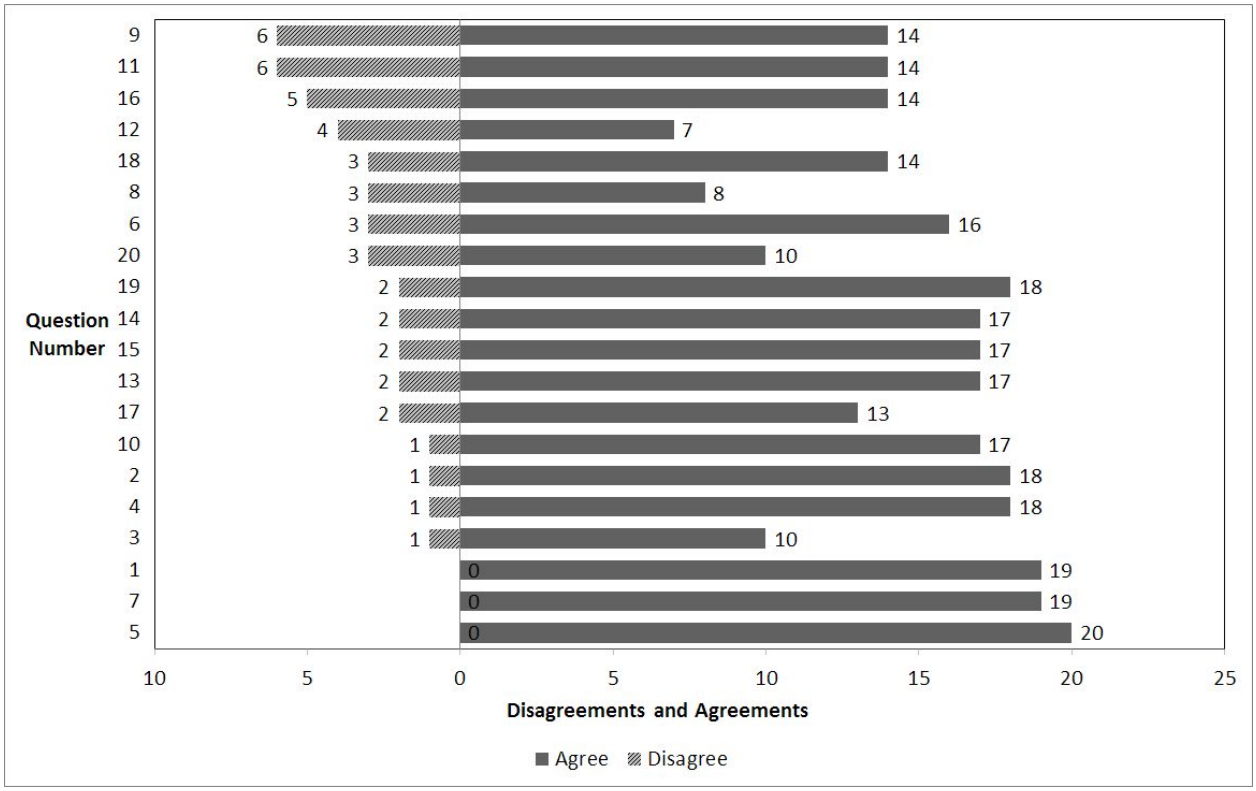


Figure 7.2G: Agreement percentage per question. Sorted by Agreement Percentage, greatest to least (best to worst).



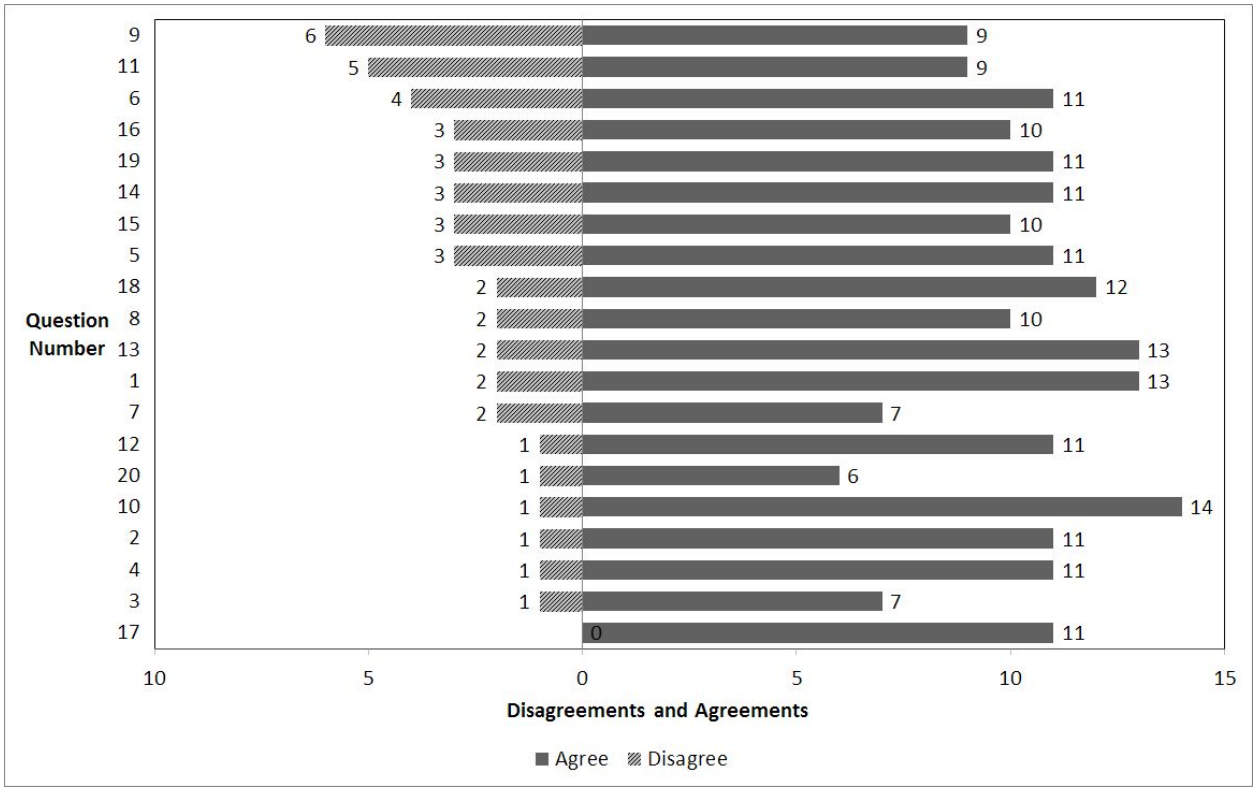
<u>Note: Figure 7.2H includes data from ONLY Just Dance Now!</u>

Figure 7.2H: Number of Agreements and Disagreements for each question. Sorted by Disagreements, greatest to least (worst to best).



<u>Note: Figure 7.2I includes data from ONLY Pokémon GO.</u>

Figure 7.2I: Number of Agreements and Disagreements for each question. Sorted by Disagreements, greatest to least (worst to best).



7.3 Validation Discussion

Validation Procedure Issues

The validation procedure was considered accurate, but could have been influenced by differences in coding between the researchers. Ideally, every single participant should have been coded by each researcher, effectively peer reviewing every single participant. It could be said that because only eight participant codings (Section 7.1) were peer reviewed, there may be inconsistency between individual researcher codings. However, it was believed the validation results were accurate given that the researchers followed the specific coding guidelines described in section 7.1. Errors caused through inconsistent codings were believed to be few, if any, and therefore negligible regarding the validation of the EEQ.

User Accounts of Problematic Questions

Of the 38 participants tested in this project, only three expressed confusion about some questions on the questionnaire. Below are the questions that were explicitly stated by participants to be confusing along with their rationale. Additionally, the participants noted that their initial interpretation of these questions impacted their answer on the EEQ.

Question 6 - "I felt that the game would have been more enjoyable without physical activity."

The participant did not understand how the game *Just Dance Now!* could be played if physical activity was not involved in the game. The user was confused as to how the game would be possible to play without dancing.

Question 17 - "I felt that the game reacted quickly to my actions."

Question 20 - "I felt in control of the game."

Both participants that mentioned these questions stated general confusion. They were not sure how to interpret what the questions were asking regarding *Pokémon GO*.

Chapter 8: Evaluating Exergame Enjoyment using the EEQ

This chapter analyzes the enjoyment data collected from the validation process (Chapter 7) for both games. The objective of enjoyment analysis was not to give a definitive score for either game that is indicative of enjoyment, but to see how the differences between these two games are reflected in the questionnaire results. It was also meant to show some different ways that EEQ results may be used.

8.1 Pokémon GO Enjoyment

At the end of the user tests, fifteen participants in total played *Pokémon GO*. The resultant enjoyment scores calculated by the EEQ scoring rules had a range from 54 to 78 (with a minimum possible score of 20 and a maximum possible score of 100). The average enjoyment score was ~67. Detailed data of EEQ results for *Pokémon GO* (including statistics per question) can be found in Appendix A.

Figure 8.1A provides the distribution of enjoyment scores for all 15 participants. Six out of fifteen participants had an enjoyment score in the range of 65 to 69. Four participants had an enjoyment score in the range of 55 to 64 and another four had an enjoyment score in the range of 70 to 79. One participant had an enjoyment score of 54: the minimum score obtained. There were no outliers. Figure 8.1B displays the cumulative probability of the enjoyment scores for *Pokémon GO*.

To analyze how *Pokémon GO* performed in each category (Immersion, Control, Reward, and Physical Activity) of the EEQ, the mean score for each question in each category was calculated. Refer to Figure 8.1C for this data. For *Pokémon GO*, the mean enjoyment score of each question of the overall EEQ was 3.3. The mean score for immersion, 2.8, was lower than the other three categories, which implies that the players did not feel immersed.

Figure 8.1A: Enjoyment score distribution for Pokémon GO

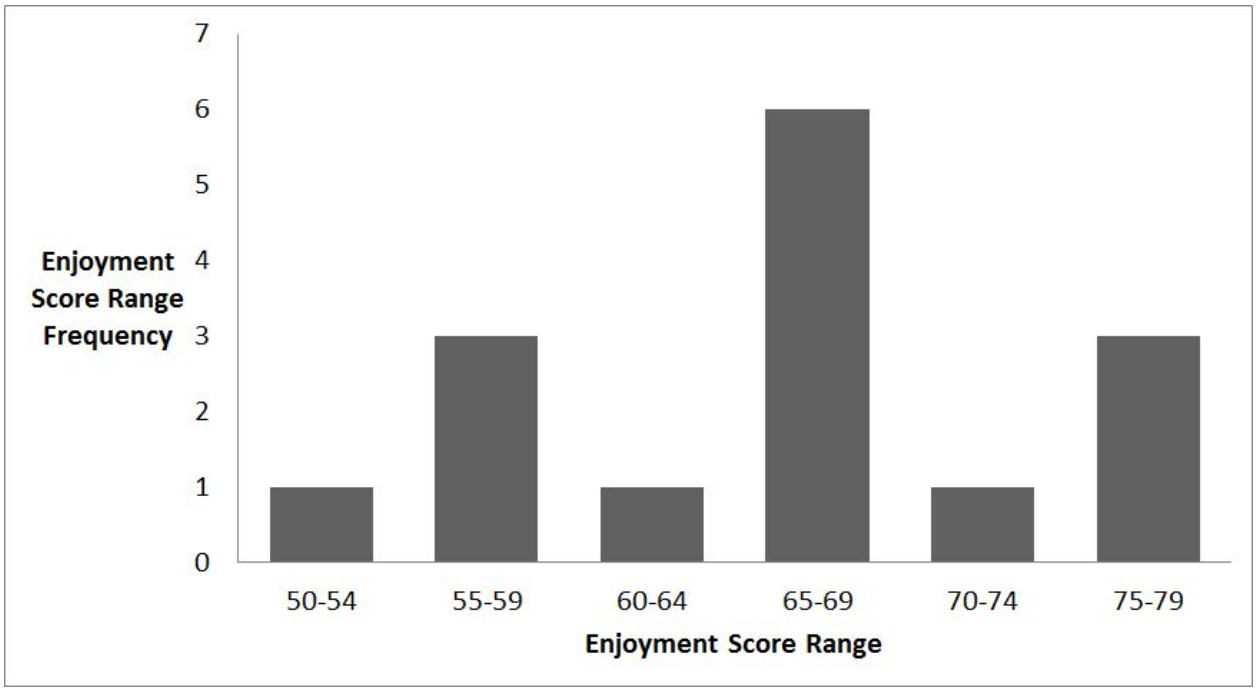


Figure 8.1B: Cumulative probability of enjoyment scores for Pokémon GO

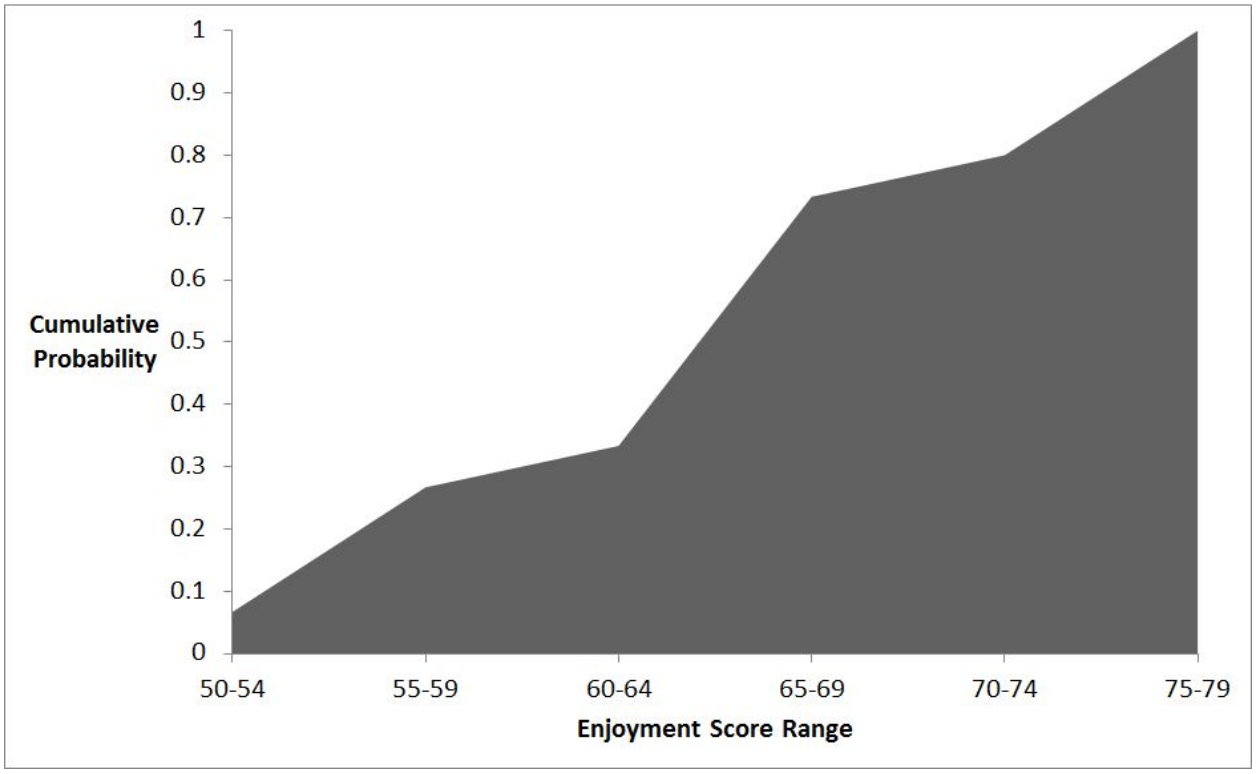


Figure 8.1C: Average score for each question and for each category for Pokémon GO

	Average score
All	3.3
Immersion	2.8
Control	3.4
Activity	3.6
Reward	3.4

8.2 Just Dance Now! Enjoyment

Twenty-three participants played *Just Dance Now.* The resultant enjoyment scores calculated by EEQ's scoring rules ranged from 59 to 96. The mean enjoyment score was ~75. Detailed data of the EEQ results for *Just Dance Now!* can be found in Appendix A.

Figure 8.2A provides the distribution of enjoyment scores from all 23 participants. Eight out of fifteen participants had an enjoyment score in the range of 70 to 74. Fewer participants fell in a higher or lower data range, which roughly follows the standard normal distribution. There was one outlier who had an enjoyment score of 96 as the maximum. This user was believed to have experienced an extreme level of enjoyment. Figure 8.2B displays the cumulative probability of the enjoyment scores for *Just Dance Now!*.

To analyze how *Just Dance Now!* performed in each category (Immersion, Control, Reward, and Physical Activity) of the EEQ, the mean score of each question of each category was calculated and displayed in Figure 8.2C. For *Just Dance Now!*, the mean score of each question of the overall EEQ was 3.7. The average score for immersion, 3.3, was much lower than the other three categories, which implies that immersion was again the weakest category for the given exergame.

Figure 8.2A: Enjoyment score distribution for Just Dance Now!

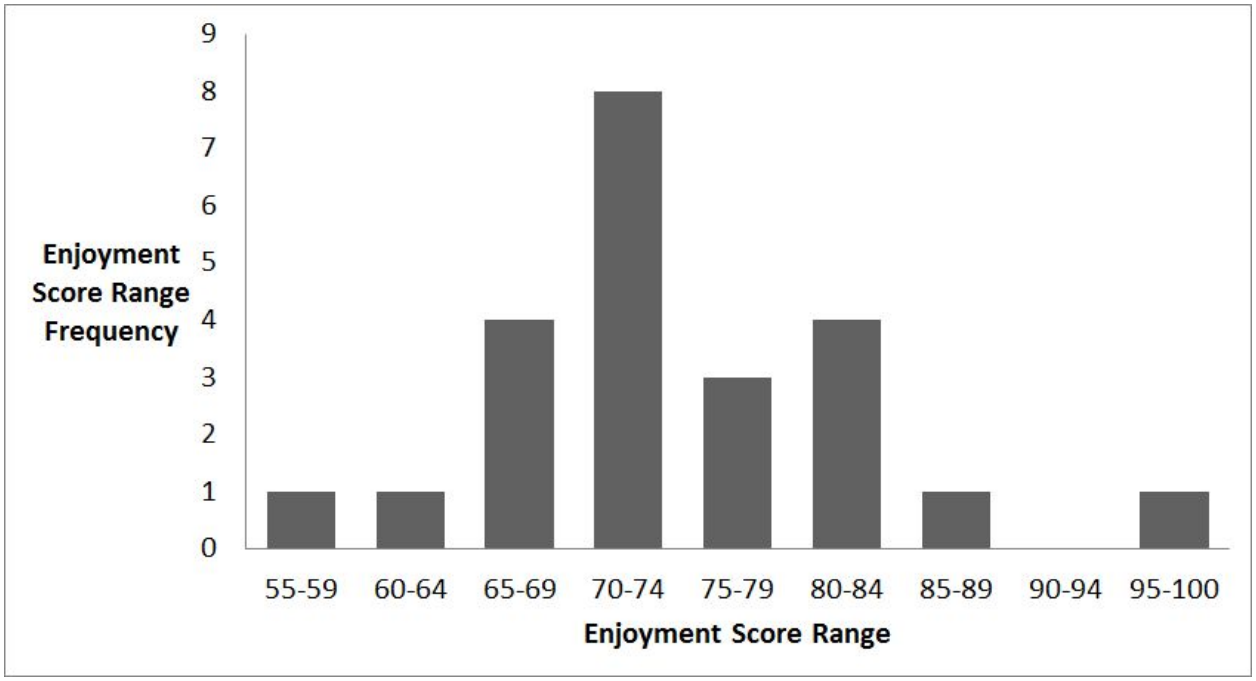


Figure 8.2B: Cumulative probability of enjoyment score for Just Dance Now!

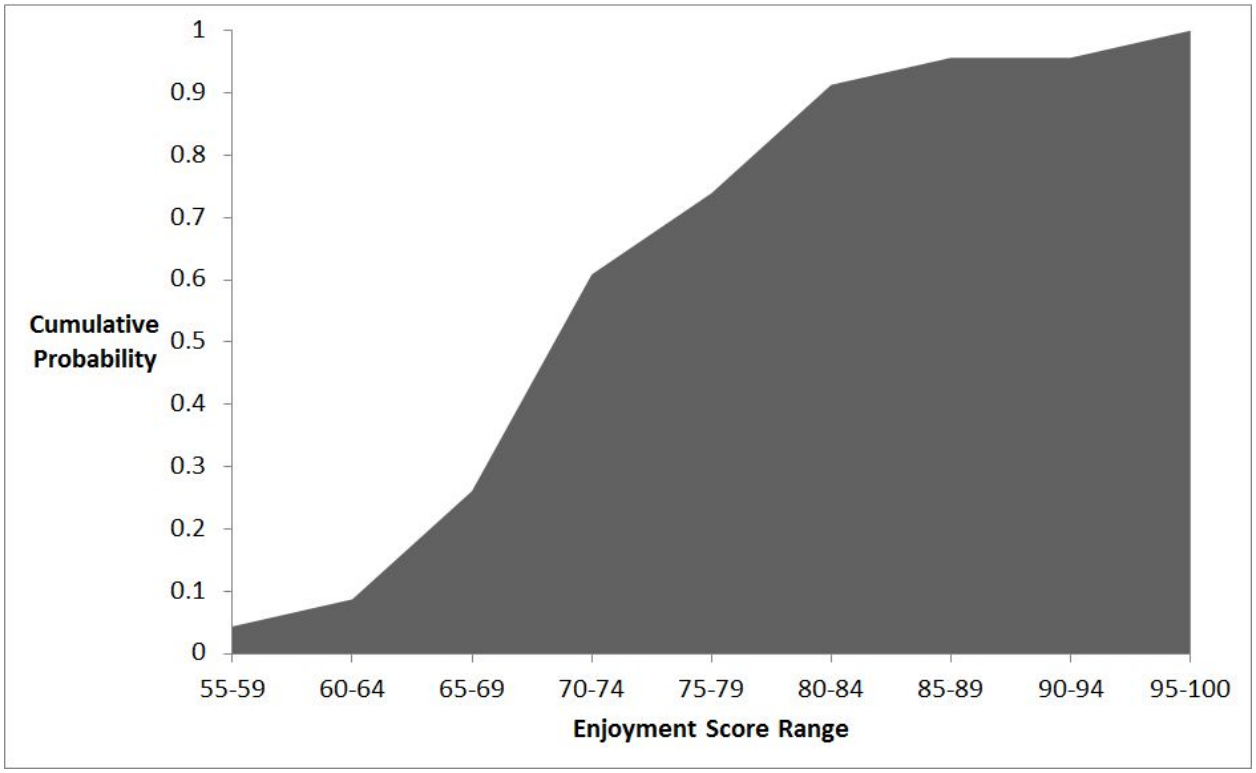


Figure 8.2C: Mean score of each question in each category for Just Dance Now!

	Average score
All	3.7
Immersion	3.3
Control	3.9
Activity	4.0
Reward	3.7

8.3 Enjoyment Discussion

Based on the results from sections 8.1 and 8.2, it appears that participants generally enjoyed *Just Dance Now!* more than *Pokémon GO*. Figure 8.3A was created by combining Figures 8.1C and 8.2C to show the comparison of the two games. As displayed in the graph, the game *Pokémon GO* not only had a lower enjoyment score, but also had a lower mean enjoyment score in every category. During user test sessions, almost all of the *Pokémon GO* participants experienced either software bugs or connection issues, which may have had an affect on their enjoyment of the game. Several users explicitly stated that these issues influenced their ability to enjoy the game. Therefore, a lower score for *Pokémon GO* was expected. However, a T-test was conducted (shown in Appendix B) and the final conclusion was that there is *not* sufficient evidence to support the claim that, on average, Just Dance Now is more enjoyable than Pokémon GO. A greater sample size is required to develop a firm conclusion.

In Figure 8.3A, the mean score for immersion was lower than the other three categories. The reason for this occurrence is uncertain, but there are several possibilities: both games are weak regarding immersion, the immersion category of EEQ questions needs revision, or players actually did not feel immersed in the games that they were playing possibly because the physical activity required in exergames might make it harder for users to get immersed compared with non-exergames.

Finally, in order to find a relationship between exergame enjoyment and the amount of physical activity a user engages in on a weekly basis, pre-questionnaire data was compared with each user's enjoyment score. This distribution is provided in Figure 8.3B. The graph suggests no correlation (line of best fit with R = 0.0007) between user enjoyment of exergames and how often the user typically exercises. Participants who exercise more often than others did not enjoy the exergames more or less than participants who engage in less physical activity.

Figure 8.3A: Comparison between mean score for questions of each section of EEQ of both games.

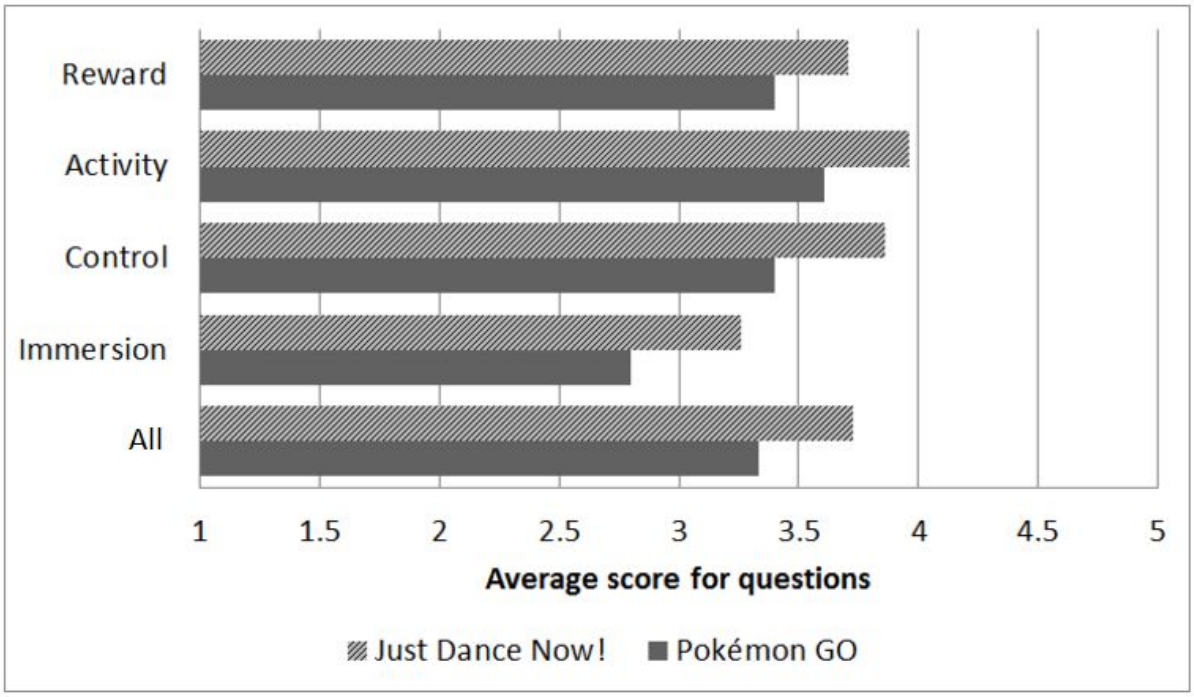


Figure 8.3B: Relationship between enjoyment score and exercise (Hours per week, both games).

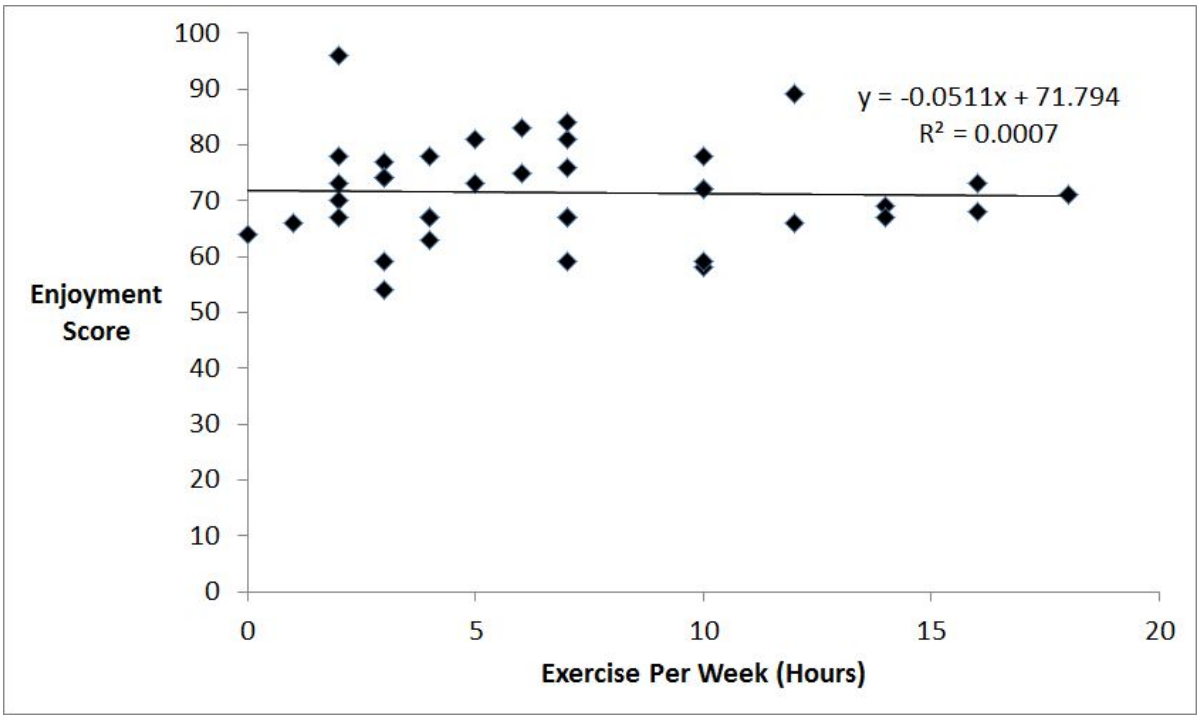
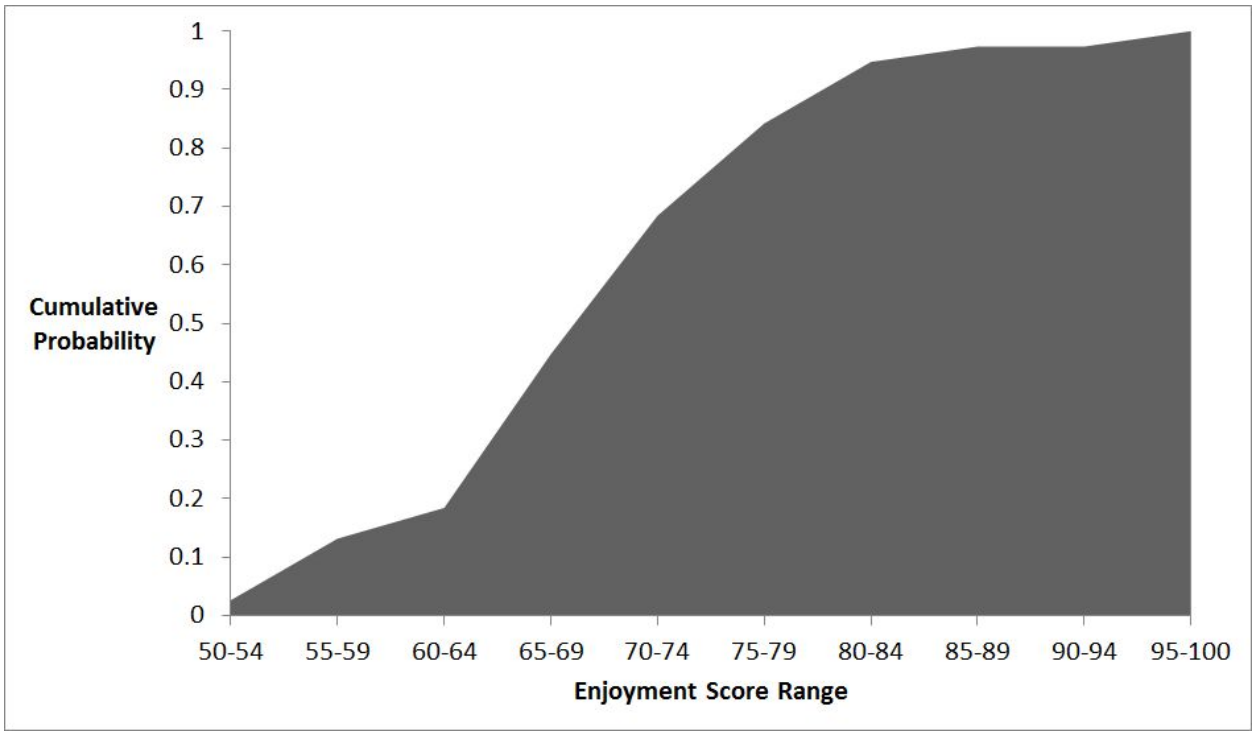


Figure 8.3C: Cumulative probability of enjoyment score for both games.



Chapter 9: Conclusion and Future Work

This chapter discusses the results of the validation process (Chapter 7). The objective of validating the Exergame Enjoyment Questionnaire was to demonstrate that the questionnaire produced reliable user exergame enjoyment data. This chapter discusses the EEQ validation process and results. Next, suggestions are given for improving the questionnaire, the methodology is discussed, and final observations/thoughts about exergame enjoyment are given. Finally, the chapter ends with some recommendations for future work.

9.1 Questionnaire Validation Results

Overall, the mean accuracy across all questions equals 84.6%, which was gathered from calculating the percentage of answers with verbal evidence to support the user's EEQ responses (Figure 7.2G). Figure 7.2A provides a visual representation, showing the number of times evidence was found for each question and Figure 7.2B shows the number of times there was no evidence for each question. It was clear that there was a positive correlation between the answers on the questions given and the responses given after gameplay. When separating the questions into their own individual categories, Intrinsically Rewarding Activity does the worst out of the four categories, giving a mean accuracy of 75.7%, which was notably lower than the rest.

As seen in 7.2C - 7.2F, every category of question has a reasonable spread. The intrinsically rewarding activity saw the largest number of no evidence questions compared to the other three categories.

When the two games were separated, many of the questions that had large amounts of contradicting evidence as well as questions that have no evidence were consistent, but there were several differences. There are several questions that were more accurate when testing one game compared to the other. One example of this is question 6, "I felt that the game would have been more enjoyable without physical activity." For *Pokémon GO*, the question has 63.6% accuracy where it has a much higher 81.3% accuracy for *Just Dance Now!*, meaning the question is 17.7% more accurate for *Just Dance Now!*. This could be because *Just Dance Now!* relies on the exercise component more heavily.

There were several factors that should be taken into account when considering the results of the study. First, there were a few variables that could have slightly affected the scores of participants. One variable was the phones given to participants for use throughout the study. The phones had Internet connectivity issues as well as software issues with the games themselves, which could have led to a comparatively negative experience.

Second, the number of participants who knew about the games was higher than expected. With the exception of one person, every participant knew of either *Pokémon GO* or *Just Dance Now!*, with the majority having heard of *Pokémon GO*. There was the possibility that users may have had preconceived opinions about certain games, leading to responses that would not change

based off of 30 minutes of gameplay. Further testing with different exergames may be necessary to determine how very popular games may have affected the results.

Third is that the number of games tested was low. As stated above, some questions were much more accurate for one game compared to the other. It is possible that this was due to a small sample size, or unknown flaws within the questionnaire that only become apparent with certain exergames.

9.2 Questionnaire Discussion

Overall, it appears that the Exergame Enjoyment Questionnaire is a pretty reliable tool for measuring exergame enjoyment. On average, each question appeared to experience agreement with their coded verbal responses roughly 85% of the time which is a fairly high percentage. However, this leaves room for improvement. It is essential to investigate which questions have the highest agreement percentages and which questions have the lowest agreement percentages.

Analysis per category of question: *Figure 7.2A* shows the number of agreements and disagreements for each question. In general, the agreement percentage appears to be a measure of how straightforward the question is or how much the players interpreted each question as it was intended. According to *Figures 7.2C-F*, the average agreement percentages for each category are:

Physical Activity: 85.6%

Immersion: 84.2%

Control: 89.7%

Intrinsic Reward: 75.7%

In particular, it appears that the EEQ's Control category is a reliable measure of how in-control the players felt. The questions appear to be easier to understand. E.g. 'I felt that it was easy to familiarize myself with the game controls' (question 7) is a statement that does not leave more room for interpretation. Alternatively, the EEQ's Intrinsic Reward category seems like an unreliable measure of how rewarding the players felt that the experience was - it could benefit from revision. Furthermore, the Physical Activity and Immersion categories each seem like a 'fairly reliable' measure of how players felt about the physical activity and immersive qualities of the games - these categories could also benefit from revision.

Most problematic questions: According to *Figure 7.2A*, the three questions with the greatest number of disagreements were questions 9, 11, and 16. These are shown below:

9. I consider playing the game 'exercise'.

11. I did not feel a desire to make progress in the game.

16. I felt a sense of accomplishment from playing the game.

These four questions could be revised (or removed) in order to improve the reliability of the questionnaire.

Question 9 asks participants if they consider the game 'exercise'. During the post-game discussion, players acknowledged that they were engaging in physical activity but did not think that it was as strenuous as a typical workout. For that reason, many participants disagreed with the statement "I consider playing the game 'exercise'" even though they agreed that the game involved a small amount of exercise. This disagreement could be solved by removing the question or by rewording it in such a way that the players do not develop a different interpretation than what was intended. E.g. The statement could read, "This game involves physical activity" or "This game involves more physical activity than a 10 minute walk.". These wordings are more straightforward and prevent confusion.

Question 11 asks participants if they did not feel a desire to make progress in the game. Some participants answered that they felt a desire to make progress in the game but called the game boring or repetitive during the post-game discussion. Others disagreed - they did not feel a desire to make progress in the game - but they expressed a desire to 'level up' in Pokémon GO or to beat their opponents in Just Dance Now. It is unclear why this question exhibited a high number of 'disagreements'. The negative wording may have confused some participants or the meaning of the word 'progress' may have been unclear. This question could be removed or revised to prevent confusion. E.g. The statement could read, "I felt a desire to accomplish tasks within the game".

Question 16 asks participants if they felt a sense of accomplishment from playing the game. Some participants answered 'neutral' to this question, but mentioned achievements during the post-game discussion. One of the post-game discussion questions asks each participant, "Did you reach any goals, accomplishments, or achievements?" This discussion question forces test participants to think about what they accomplished. It is possible that they actually had no feeling of accomplishment but implied that they did by answering the discussion question. This discussion question could be replaced with something along the lines of, "How did you feel about leveling up in Pokémon GO?" or "How did you feel about competing with the other dancers?" These types of questions give participants the opportunity to say if they accomplished anything, but do not specifically lead them to agree with the statement that they achieved something. If the questions asked in the game were directly related to the game, it would be easier to code. The reliability of question 16 is unclear at this time.

The questionnaire also contains questions that had a high percentage of 'no-evidence' conclusions. This issue is discussed in Section 9.3.2, which reflects upon the coding process.

9.3 Methodology Discussion

The sections below describe how the testing and coding steps could be improved.

9.3.1 *Improving Testing*

The testing procedure was very standardized: created with the intention of easy replication. There were two main stages of testing: the focus groups and the validation.

In the focus groups phase of testing, the questionnaire would be changed for every session, with the intention of making the questionnaire as accurate and simple as possible before beginning the validation process. Each subsequent version of the questionnaire would revise issues that participants had. However, it is clear now that the questionnaire would have benefited from a longer and more thorough focus group phase. Many of the problems and inconsistent questions might have been prevented if more time was spent before starting the validation step.

Another factor that was apparent in the testing stages was that the actual phones given to participants in order to play the exergames were of low quality. These phones had poor WiFi connectivity and also had a habit of crashing during gameplay. These problems could have affected the resulting enjoyment scores. In the future, the mobile phones should be determined as stable and effective before being given to participants.

9.3.2 *Improving Coding*

Figure 7.2B shows the number of 'No evidence' conclusions for each question after conducting the coding process. The four questions with the greatest number of 'No evidence' were questions 3, 8, 12, and 20. These are shown below:

> 3. I felt like I lost track of time while playing.
>
> 8. I felt emotionally attached to the game.
>
> 12. I felt a strong sense of being in the world of the game to the point that I was unaware of my surroundings.
>
> 20. I felt in control of the game.

The coding process as a whole could be improved by adding questions to the post-game discussion that target these EEQ questions. In particular, it seems that the immersion category is being somewhat ignored during the post-game discussions. Participants could be asked to describe their level of focus while playing the game.

9.4 Final Observations

This section describes observations that the test proctors made during testing. Unlike the data provided in Chapters 7 and 8, these observations relate to unquantifiable information which cannot be represented graphically.

9.4.1 Validity of the Exergame Enjoyment Questionnaire

Not only did the EEQ have a high average agreement percentage per question (of ~85%), but the highest scores on the EEQ seemed consistent with the players that expressed the most joy during the post-game conversation. The highest scores all came from participants who had played Just Dance Now. These top scoring participants seemed to love the game. For that reason, the EEQ seems to be a reliable tool for identifying those players who enjoy the exergame the most. Unfortunately, there is no obvious correlation between the medium or low-scoring participants and the joy that they expressed afterwards.

9.4.2 Reliability of enjoyment data

The overall sample size of 38 players has been considered enough people for the validation statistics to be considered 'statistically significant'. However, when split between two games, this is not enough people for the enjoyment statistics to be considered 'statistically significant'. A total sample size of 60 players or more (more than 30 per game) would have allowed more reliable conclusions to be made. In particular, the enjoyment data is also affected by some users' past experiences with the games. Multiple Pokémon GO participants stated that they had played the game a lot prior to the testing session. It is possible that they enjoyed the game less because they had already played it - some former Pokémon GO players even admitted that they had become bored with the game. If the EEQ was tested for different exergames that people were less familiar with, then the enjoyment data may become more reliable.

9.4.3 Effect of playing Just Dance Now! in a party setting.

The average enjoyment score for players who played Just Dance Now with other participants was 0.73. The average enjoyment score for players who played Just Dance Now alone was 0.79. The evidence suggests that it is more fun to play the game alone, however, these sample sizes are especially low with less than ten players playing solo. Furthermore, the top two highest scoring participants also happened to play alone; this could have contributed to a slightly misleading result. Understanding the relationship between enjoyment and number of players will require additional testing. From the post-game discussions, it is clear that some players feel uncomfortable dancing alone in front of the test proctors and would prefer to have company. However, other players have said that they are glad they played alone for fear of embarrassing themselves in front of other players. It is unclear if enjoyment and number of players are strongly related.

9.5 Recommendations for Future Work

While the analysis and results have finished, there is still quite a bit of room for further development for the EEQ, including but not limited to: fixing and changing specific questions, changing testing environments, testing different and a larger variety of games as well as testing with a larger user sample size.

After the validation process, while many of the questions tested in the questionnaire were accurate in measuring user enjoyment of exergames, there were some questions that could be improved, either because there were too many conflicts within the discussion or there was not enough evidence to analyze the question. These questions should be modified in order to produce a more accurate association between the discussion responses and the EEQ answers.

Among certain problems with the project methodology, the user sample size was small. A sample size many times greater than that used for this project would help to further analyze the data gathered in this study and yield statistically significant conclusions. Testing a greater number of games in different situations with a more diverse set of users would help test the EEQ more thoroughly.

Another area of improvement could be pre-questionnaires. The pre-questionnaire could have more questions relating directly to the test. For example, questions that relate directly to the testing such as, “Do you think you’d feel embarrassed playing an iPhone game alone/with a group?”

With a larger group of participants, it is likely that much of the variation within the EEQ would be removed, giving a more clear and concise picture of the accuracy of the EEQ. The removal or modification of problematic questions would be expected to increase the accuracy of the EEQ in terms of measuring user enjoyment of exergames.

References

1. Beadell, Stephanie. "Surveys 101: A Simple Guide to Asking Effective Questions." Zapier - The Ultimate Guide to Forms and Surveys. N.p., n.d. Web. 11 Sept. 2016.
2. Brockmyer, J. H., Christine, F. M., Curtiss, K. A., McBroom E., Burkhart, K. M., Pidruzny, J. N., *The Development of the Game Engagement Questionnaire: A Measure of Engagement in Video Game-playing*. *The Development of the Game Engagement Questionnaire: A Measure of Engagement in Video Game Playing: Response to Reviews*. Journal of Experimental Social Psychology 45 (2009) 624–634, 12 Mar. 2009. Web. 10 Sept. 2016.
3. "Control." Def. 1. *Oxford Dictionaries*. N.p., n.d. Web. 5 Oct. 2016.
4. Cox, A. L., Jennett, C., Paul C., Dhoparee, S., Epps, A., Tijs, T., Walton, A., *Measuring and Defining the Experience of Immersion in Games.* International Journal of Human-Computer Studies 66.9 (2008): 641+. Web. 5 Sept. 2016.
5. "exergame." YourDictionary, n.d. Web. 5 October 2016.
6. "Enjoyment." Def. 1. *Merriam-Webster*. N.p., n.d. Web. 5 Oct. 2016.
7. "Finding the Correct Survey Length - FluidSurveys." FluidSurveys. Fluidsurvey's Team, 7 Aug. 2013. Web. 02 Oct. 2016.
8. IJsselsteijn, W. A. *GAME EXPERIENCE QUESTIONNAIRE*. Technische Universiteit Eindhoven, n.d. Web. 1 Sept. 2016.
9. "Immersion." Def. 3. *Dictionary.com*. N.p., n.d. Web. 5 Oct. 2016.
10. "Intrinsic Reward." *Business Dictionary.* N.p., n.d. Web. 5 Oct. 2016.
11. *Just Dance Now*. 1.5.1. Ubisoft. 2014. Video Game.
12. *Pokémon GO*. 0.39.1. Niantic Labs. 2016. Video Game.
13. Mcveigh, Joanne A., Elisabeth A. H. Winkler, Erin K. Howie, Mark S. Tremblay, Anne Smith, Rebecca A. Abbott, Peter R. Eastwood, Genevieve N. Healy, and Leon M. Straker. "Objectively Measured Patterns of Sedentary Time and Physical Activity in Young Adults of the Raine Study Cohort." *Int J Behav Nutr Phys Act International Journal of Behavioral Nutrition and Physical Activity* 13.1 (2016): n. pag. Web. 28 Sept. 2016.

14. “Measuring Enjoyment of Physical Activity in Children: Validation of the Physical Activity Enjoyment Scale.” *Journal of applied sport psychology* 21.S1 (2009): S116–S129. *PMC*. Web. 5 Oct. 2016.
15. Mullen, Sean P et al. “Measuring Enjoyment of Physical Activity in Older Adults: Invariance of the Physical Activity Enjoyment Scale (paces) across Groups and Time.” *The International Journal of Behavioral Nutrition and Physical Activity* 8 (2011): 103. *PMC*. Web. 5 Oct. 2016.
16. Nordin, Imran A., Alena Denisova, and Paul Cairns. "Too Many Questionnaires: Measuring Player Experience Whilst Playing Digital Games." *Too Many Questionnaires: Measuring Player Experience Whilst Playing Digital*
17. OUM SAOKOSAL, Chief of Faculty of Computer Science Follow. "Measuring And Defining The Experience Of Immersion In Games." *Measuring And Defining The Experience Of Immersion In Games*. N.p., 28 Sept. 2009. Web. 05 Sept. 2016.
18. Ryan, By Scott RigbyRichard. "Introduction." Gamasutra. N.p., n.d. Web. 11 Sept. 2016. <http://www.gamasutra.com/view/feature/130155/rethinking_carrots_a_new_method_.php?print=1>.
19. Staiano, Amanda E., and Sandra L. Calvert. "Exergames for Physical Education Courses: Physical, Social, and Cognitive Benefits." *Child Development Perspectives*. The Society for Research in Child Development, 10 May 2011. Web. 28 Sept. 2016.
20. Swatman, Rachel. "Pokémon Go Catches Five New World Records." *Pokémon Go Catches Five New World Records*. Guinness World Records, 10 Aug. 2016. Web. 08 Oct. 2016.

Appendices

Appendix A

Both Games Exergame Enjoyment Questionnaire Data

Category	Question	Mean	Median	Min	Max	Std Dev	Variance
ACTIVITY	**10**	4.32	5	2	5	0.86	0.74
ACTIVITY	**6**	4.18	4	1	5	0.91	0.83
ACTIVITY	**13**	3.84	4	1	5	1.06	1.13
ACTIVITY	1	3.82	4	2	5	0.72	0.52
ACTIVITY	2	3.79	4	2	5	0.77	0.59
ACTIVITY	14	3.79	4	1	5	0.86	0.75
ACTIVITY	**19**	3.63	4	2	5	1.06	1.13
ACTIVITY	9	3.21	3	1	5	1.08	1.17
CONTROL	7	3.97	4	2	5	0.87	0.76
CONTROL	20	3.84	4	2	5	0.78	0.61
CONTROL	**4**	3.82	4	1	5	1.05	1.10
CONTROL	17	3.08	3	1	5	1.16	1.34
IMMERSION	5	3.82	4	1	5	0.79	0.62
IMMERSION	3	3.53	4	1	5	1.07	1.14
IMMERSION	**18**	3.37	3	1	5	0.93	0.86
IMMERSION	8	2.42	2	1	5	0.94	0.88
IMMERSION	12	2.26	2	1	5	0.99	0.98
REWARD	15	3.84	4	1	5	0.87	0.76
REWARD	**11**	3.47	4	1	5	1.04	1.09
REWARD	16	3.45	4	2	5	0.94	0.88

Pokémon GO Exergame Enjoyment Questionnaire Data

Category	Question	Mean	Median	Min	Max	Std Dev	Variance
ACTIVITY	1	3.47	3	2	5	0.72	0.52
ACTIVITY	2	3.53	4	2	5	0.81	0.65
IMMERSION	3	3.13	3	1	5	1.20	1.45
CONTROL	4	3.47	4	1	5	1.15	1.32
IMMERSION	5	3.53	4	2	4	0.72	0.52
ACTIVITY	6	4.13	4	2	5	0.72	0.52
CONTROL	7	4.00	4	2	5	0.89	0.80
IMMERSION	8	2.07	2	1	4	0.77	0.60
ACTIVITY	9	2.73	3	1	4	0.93	0.86
ACTIVITY	10	4.53	5	2	5	0.81	0.65
REWARD	11	3.67	4	2	5	0.70	0.49
IMMERSION	12	2.20	2	1	4	0.91	0.83
ACTIVITY	13	3.87	4	1	5	1.09	1.18
ACTIVITY	14	3.60	4	1	5	0.95	0.91
REWARD	15	3.53	4	2	5	0.81	0.65
REWARD	16	3.00	3	2	4	0.73	0.53
CONTROL	17	2.40	2	1	4	1.02	1.04
IMMERSION	18	3.07	3	1	4	0.85	0.73
ACTIVITY	19	3.00	3	2	5	0.97	0.93
CONTROL	20	3.73	4	2	5	0.85	0.73
	Enjoyment Score	**66.67**	**67**	**54**	**78**	**7.01**	**49.16**

Just Dance Now! Exergame Enjoyment Questionnaire Data

Category	Question	Mean	Median	Min	Max	Std Dev	Variance
ACTIVITY	1	4.04	4	3	5	0.62	0.39
ACTIVITY	2	3.96	4	3	5	0.69	0.48
IMMERSION	3	3.78	4	2	5	0.88	0.78
CONTROL	**4**	4.04	4	1	5	0.91	0.82
IMMERSION	5	4.00	4	1	5	0.78	0.61
ACTIVITY	**6**	4.22	4	1	5	1.02	1.04
CONTROL	7	3.96	4	2	5	0.86	0.74
IMMERSION	8	2.65	3	1	5	0.96	0.92
ACTIVITY	9	3.52	4	2	5	1.06	1.12
ACTIVITY	**10**	4.17	4	2	5	0.87	0.75
REWARD	**11**	3.35	4	1	5	1.20	1.44
IMMERSION	12	2.30	2	1	5	1.04	1.08
ACTIVITY	**13**	3.83	4	1	5	1.05	1.10
ACTIVITY	14	3.91	4	2	5	0.78	0.60
REWARD	15	4.04	4	1	5	0.86	0.74
REWARD	16	3.74	4	2	5	0.94	0.89
CONTROL	17	3.52	4	1	5	1.02	1.03
IMMERSION	**18**	3.57	4	1	5	0.92	0.85
ACTIVITY	**19**	4.04	4	2	5	0.91	0.82
CONTROL	20	3.91	4	2	5	0.72	0.51
	Enjoyment Score	**74.57**	**73**	**59**	**96**	**8.26**	**68.25**

Appendix B: T-test for Mean Enjoyment Scores

Pokémon GO	PGO^2	Just Dance Now	JDN^2
66	4356	75	5625
77	5929	72	5184
76	5776	67	4489
58	3364	72	5184
59	3481	74	5476
54	2916	81	6561
66	4356	84	7056
63	3969	89	7921
78	6084	64	4096
69	4761	83	6889
68	4624	71	5041
59	3481	78	6084
67	4489	73	5329
67	4489	78	6084
73	5329	70	4900
		59	3481
		96	9216
		67	4489
		67	4489
		73	5329
		81	6561
		74	5476
		67	4489

Hypothesis: The mean enjoyment score for Just Dance Now is greater than the mean enjoyment score for Pokémon GO.

$\left(\sum PGO\right)^2 = 1,000,000$ and $\left(\sum JDN\right)^2 = 2,941,225$

$\mu_{PGO} = 66.7$ and $\mu_{JDN} = 74.6$

$\sum PGO^2 = 67,404$ and $\sum JDN^2 = 129,449$

$n_{PGO} = 15$ and $n_{JDN} = 23$

$$A = \sum PGO^2 - \frac{\left(\sum PGO\right)^2}{n_{PGO}} \text{ and } B = \sum JDN^2 - \frac{\left(\sum JDN\right)^2}{n_{JDN}}$$

$$t = \frac{\mu_{PGO} - \mu_{JDN}}{\sqrt{\left[\frac{A+B}{n_{PGO} + n_{JDN} - 2}\right] \cdot \left[\frac{1}{n_{PGO}} + \frac{1}{n_{JDN}}\right]}} = -0.07377$$

Degrees of freedom $= n_{PGO} - 1 + n_{JDN} - 1 = 36$

At an alpha level of 5% and 36 degrees of freedom, $t_a = 2.034 > 0.07377$

Conclusion: There is insufficient evidence to support the claim that the mean enjoyment score for Just Dance Now is greater than the mean enjoyment score for Pokémon GO.