Dominik Christian Widnig, BSc

Evaluation of User Experience in a Location-Based Mobile Role-Playing Game

Master's Thesis

to achieve the university degree of
Diplom-Ingenieur

Master's degree programme:
Software Engineering and Management

submitted to
Graz University of Technology

Supervisor
Ass.-Prof. Dr. Viktoria Pammer-Schindler

Knowledge Technology Institute
Head: Univ.-Prof., Dr. Stefanie Lindstaedt

Graz, January 2017

Statutory Declaration

I declare that I have authored this thesis independently, that I have not used other than the declared sources/resources, and that I have explicitly marked all material which has been quoted either literally or by content from the used sources.

Graz, _____ _____

Date Signature

Eidesstattliche Erklärung[1]

Ich erkläre an Eides statt, dass ich die vorliegende Arbeit selbstständig verfasst, andere als die angegebenen Quellen/Hilfsmittel nicht benutzt, und die den benutzten Quellen wörtlich und inhaltlich entnommenen Stellen als solche kenntlich gemacht habe.

Graz, am _____ _____

Datum Unterschrift

[1] Beschluss der Curricula-Kommission für Bachelor-, Master- und Diplomstudien vom 10.11.2008; Genehmigung des Senates am 1.12.2008

Abstract

Location-based games are currently more popular than ever for the general public. Games, such as *Geocaching*, *Ingress* and *Pokemon Go* have created a high demand in the app market and established themselves in a major category in the mobile gaming sector. Since location-based games are reliant on mobile sensors, battery life, cellular data connections and even environmental conditions, many problems can rise up while playing the game and hence, can reduce user experience and player enjoyment. The aim of this thesis is to improve the gaming experience of location-based games, which use map information to place virtual content at appropriate physical locations, with the assistance of an user-centered design approach. Therefore, a game named *Geo Heroes* was designed and implemented in order to evaluate it with existing quantitative and qualitative methods from research. The game was assessed in an empirical study with nine participants including a game-play session of about one hour. Participants were divided into an experimental and control group to author disparities in the implemented content placement algorithms. An already established questionnaire for traditional computer games, and one created by the author based on existing research in location-based games, were used to measure common factors in gaming experience. Additionally, participants sent log data with their current emotions during game-play after various interactions with game objects. Different outcome scenarios of interactions were considered to ensure a better analysis. Furthermore, an open group discussion was held to gather qualitative information from participants to reveal still undiscovered issues and to provide evidence from results of conducted quantitative methods. Results have shown that the questionnaire for location-based games is a useful tool to measure player enjoyment. In combination with the tracked emotions and a group interview, relevant information can be obtained in order to improve game design and mechanics.

Kurzfassung

In der Gesellschaft sind Location-based Games momentan beliebter denn je. Spiele wie *Geocaching*, *Ingress* und *Pokemon Go* erzielten eine hohe Nachfrage in den App Märkten und etablierten sich dadurch zu einer wichtigen Kategorie im Mobile-Gaming Sektor. Da Location-based Games abhängig von mobilen Sensoren, Akkulaufzeit, Datenverbindung und sogar Umweltbedingungen sind, können viele Probleme während des Spielens auftreten, welche wiederum Auswirkungen auf die User Experience verursachen. Ziel dieser These ist es, mithilfe von User Centered Design, die User Experience von Location-based Games zu verbessern, welche Karteninformationen verwerten um ihre virtuellen Inhalte auf reale Orte (Positionen) abzustimmen und anzupassen. Aufgrund dessen wurde ein Spiel, namens *Geo Heroes* entwickelt, welches mit bereits existierenden empirischen, quantitativen und qualitativen Forschungsmethoden, inklusive einer einstündigen Spielzeit mit neun Teilnehmern, evaluiert wurde. Diese wurden in eine Experimentier- und Kontrollgruppe eingeteilt um Unterschiede zwischen den Algorithmen für Spielinhaltsplatzierungen aufzuzeigen. Verwendet wurde ein etablierter Fragebogen für traditionelle Computerspiele und ein selbst erstellter Fragenkatalog für location-based Games, um User Experience Faktoren im Bereich Gaming zu messen. Es wurden Logdaten aufgezeichnet, welche die Emotionen von Teilnehmern während der Spielzeit nach diversen Interaktionen beinhalten. Ein Gruppeninterview mit allen Teilnehmern diente dem Erhalt weiterer qualitativer Informationen. Dieses zeigte unentdeckte Probleme auf, war aber auch förderlich für die Bestätigung von erzielten Ergebnissen in den quantitativen Analysen. Dieses Resultat veranschaulichte, dass Evaluierungsmethoden die auf location-based Games angepasst wurden ein nützliches Werkzeug zur Messung von User Experience sind. In Kombination mit den aufgezeichneten Emotionen und einem Gruppeninterview, können relevante Informationen extrahiert werden, um damit Spieldesign und -mechaniken von solchen Spielen zu verbessern.

Contents

List of Figures

List of Tables

List of Listings

1 Introduction

The main purpose of this thesis is to design and evaluate a location-based game played on mobile devices called *Geo Heroes* in order to be able to improve it. A first prototype was already developed in a previous project by the author. Further design and performance improvements are included in the course of this thesis. Related work regarding the evaluation of user experience in traditional and location-based games is conducted. The thesis focuses especially on the terms of immersion and game flow in pervasive games, an umbrella term for location-based games. With the current implementation state it is difficult for developers and designers to decide which elements in the game should be continued and focused on and which ones should be abandoned without feedback of potential users. Hence, development can go in the wrong direction and additional costs may arise due to false design decisions. A user-centered design approach can therefore help to overcome such risks and can provide valuable feedback in order to improve the game. Gained information can be used to integrate features which will increase user experience and player enjoyment. Although location-based games are not new in the research area, economy-wise they are still a niche product in the mass market of mobile games. Currently there are a few big players in the market such as *Geocaching*[1], *Ingress*[2] and *Pokemon Go*[3], the latter being released during the finalization of this thesis. *Geo Heroes*, being a location-based role-playing game in a medieval-fantasy setting, could also establish itself as another representative in this market. Since the development of such games is more complex compared to traditional mobile games, due to the usage of location-aware technologies such as GPS and cellular data connection, player enjoyment can be drastically

[1] https://www.geocaching.com/
[2] https://www.ingress.com/
[3] http://www.pokemongo.com/

decreased in case of failures of mentioned technologies. As a result, players will stop playing and probably never return to the game. In order to prevent such a scenario, a user-centered design approach during the development phase of this type of games could uncover potential design and technical problems. Hence, this thesis conducts an empirical study and compares the usefulness of existing quantitative and qualitative methods from reviewed literature to gain results and to improve *Geo Heroes* for a next iteration in the design process. Additionally, this thesis covers the use of an algorithm based on map information, which is responsible for placing virtual content to appropriate physical locations in the real world. Since location-based games enable physical activities for players, a comparison between players with and without gaming experience in traditional computer games is also conducted.
The thesis is structured in the following chapters. First, a literature review of related work is conducted. Second, a research question with hypotheses is constructed. In the subsequent chapters, the game itself is presented to give the reader an overview on game content, since some of the contents strongly relate to the chosen evaluation method which is conducted as an empirical study with nine participants. Gained results of the evaluation are presented and discussed. Finally, possible improvements for *Geo Heroes* are mentioned and insights are shared for future work in this area.

2 Background and Related Work

The chosen literature consists of topics related to different evaluation methods for measuring user experience. Evaluation techniques for traditional and pervasive games are analyzed in the course of this thesis and each type is described in its own section. Since this thesis also covers the design and implementation aspect of a location-based game, an extra section is introduced describing this subject. Before the mentioned sections starts, relevant definitions in this research area are described.

2.1 Conceptual Background of Presence, Immersion and Game Flow

There are always the same recurring terms when reviewing literature describing user experience in games. These terms are used by researchers for a common way of expression and a consistent scale. In order to provide a common knowledge base, relevant terms for this thesis are explained in detail.

Presence

Presence is a term that was mainly used in the past when researchers started to investigate virtual environments (VE). Participants had to wear a helmet-mounted display in order to be in such a VE and *Presence* described their current emotional state. Presence can therefore be seen as the psychological sense of being in a VE (Jennett et al., 2008). Relevant factors to measure Presence are control, sensory, distraction and realism, which were further investigated by Witmer and Singer (1998) in constructed questionnaires.

Immersion

As time passed on, new terms such as immersion were introduced when talking about gaming experience (GX). Immersion or being immersed is often used by participants when they try to explain their current situation in a virtual world. A virtual world does not mean to be a VE, thus immersion can be reached also in computer games. Witmer and Singer (1998) describe immersion "...as a psychological state characterized by perceiving oneself to be enveloped by, included in, and interacting with an environment that provides a continuous stream of stimuli and experiences." Brown and Cairns (2004) state that total immersion is presence. However, literature is still debating about the relationships of these terms. For example, Jennett et al. (2008) argue that Presence is a state of mind and immersion an experience in time. Immersion can therefore be seen as the engagement in a computer game and the key to a good gaming experience. Sanders and Cairns (2010) state that "...immersion in this sense is a property of the system that allows a person to feel presence." As an example, the game *Tetris* probably will not trigger presence, since nobody would feel to be in the world of fallen blocks. However, *Tetris* was a popular game in the past and the game triggers a sense of time loss and pure concentration within the game which leads to forget all surrounding factors in the real world. This phenomenon describes immersion at its best. Even if players do not feel presence, they can still be immersed. Additionally it can be said, that presence is possible even if not being immersed. Imagine a VE where participants have to do boring and monotonous tasks. They would still feel to be empathized in a virtual world even if physically situated in another, but properties which lead to immersion would not exist. Immersion is also described in Cox et al. (2012) as "...the sense of being absorbed in a game to the exclusion of all else outside of the game." Immersion is also an important component to bring players back to a game (Huhtala et al., 2012).

Game Flow

The term *Game Flow* was introduced by Sweetser and Wyeth (2005). It is based on the concept of *Flow* which was initially described by Csikszentmihalyi (1990) and can be seen as the optimal experience in enjoyment independent from culture and type of activity. Hence, *Game Flow* is the extension and adaption to the computer gaming genre and emerges when

the balance between player skill and challenge is given, which results in an intrinsic rewarding activity (Brockmyer et al., 2009). Cox et al. (2012) describe the concept of flow in terms that are used in the field of sports. Athletes are in flow when they are "in the zone" of a game. Therefore, the authors argue that flow as well as game flow are either experienced or not. Athletes can not be "a bit in the zone", and therefore the concept of flow can be seen as a binary decision. However, the relationships between flow, immersion and presence is still not clear in the literature of gaming (Jennett et al., 2008)(Huhtala et al., 2012).

2.2 Evaluation of Traditional Computer Games

While there is a number of standardized frameworks in measuring user experience in the field of applications, a common framework to determine which method should be used for evaluating various interaction concepts in games does not exist (Bernhaupt et al., 2007). However, there are many publications describing the evaluation in different areas of gaming.
Brockmyer et al. (2009) introduced a questionnaire for measuring the potential for becoming engaged in computer games especially in the violence genre. The term engagement is described by the authors as a generic indicator of game involvement. The resulting *Game Engagement Questionnaire* consists of 19 items, containing questions which can be categorized in the terms of flow, immersion, presence and psychological absorption. Gee (2003) talks about the effect of learning in good games. These games motivate individuals by actively involving players in scenarios.

> *Motivation is the most important factor that drives learning. When motivation dies, learning dies and playing stops. - Gee (2003)*

In games, players are forced to create solutions instead of just being the inactive part as it is often in school. Players are rather producers than consumers. When being a producer, people have to come up with ideas on how to solve a problem.
Games confront players with similar types of problems. The player is forced to think of solutions for different problems. These solutions then result

in certain skills achieved through a routinized mastery. This behavior is repeated in games with new types of problems to find new ways of solving it and to create a new higher-order routinized mastery skill. This so called "cycle of expertise" is the foundation for creating expertise in all areas. Good games should therefore be designed to allow players to recreate themselves in a new world and to gain new knowledge from profound learning. Huhtala et al. (2012) measured the immersion rate by using the questionnaire presented in Jennett et al. (2008). Each of the 31 questions is assigned to one of the six different categories: *Basic attention, temporal dissociation, transportation, challenge, emotional involvement and enjoyment*. The authors evaluated two mobile games. One game called *Fozzles* was still under development and close to a release, the other game was *Angry Birds* for benchmark. Nine participants were recruited for the evaluation experiment, consisting of two sessions of game play. In session one, the first version of *Fozzles* was evaluated with the questionnaire. One month later the improved version of the same game was evaluated again. Statistical significance was measured by applying pairwise t-tests. However, the evaluation results of the two versions were a surprise. Although the developers had improved the game based on the results of the questionnaire of the first version, the new version had a lower immersion rate than the first version. The authors are therefore skeptical about techniques for measuring immersion such as this questionnaire especially when it comes to a fine tuning process. They also mention the boredom factor of a game played the second time and the experience factor in a game such as *Angry Birds* which could bias the results of measuring immersion. However, due to the division of immersion categories in the questionnaire, the differences in each category were high. The questionnaire can therefore help to analyze and measure different types of immersion. It is then possible to focus on the lowest scoring areas to improve the overall immersion result. Due to the fact that games are nonlinear and interactive, Qin et al. (2007) developed a questionnaire for measuring player immersion in the computer game narrative. In their original questionnaire 30 items were used to measure general dimensions of immersion including story comprehension and emotion. After a field experiment with over 300 participants on a website, factor analyses and a reliability test of the results, the authors could extract six dimensions of immersion. The answering of the questionnaire was conducted under the condition that participants had to imagine a familiar game with a story. Fur-

ther work was suggested and done by the same authors two years later (Qin et al., 2009). After applying *Exploratory Factor Analysis (EFA)*, the original six factor-model was modified. The optimum number of factors was achieved by using *Confirmatory Factor Analysis (CFA)*. With the *CFA* the number of components was increased by one, which resulted in a seven-factor model. The authors therefore added and named the seventh dimension *Familiarity*. Cross-validation ensures correctness of applied analysis. A year later Qin et al. (2010) used this questionnaire to evaluate player immersion for different game difficulties. The study was performed with 48 participants playing the same game with different difficulty settings. The provided difficulties were applied per level and not for the whole game. Two different approaches were used. First, the direction of game difficulty was tested. There were three different directions: Continuous, up-and-down and down-and-up. The down-and-up setting provides a higher difficulty at the beginning and at the end of the level, whereas the up-and-down setting uses a higher difficulty in the middle and at the end of the level. The second approach was to change the overall value of difficulty for each level. Three options were tested in the scenarios: slow, medium and a fast increase between two connecting levels. The results of the field experiment revealed that player immersion is highest when using the up-and-down option in combination with a medium rate of difficulty change. Since the difficulty of a game relates to the level of challenge, Cox et al. (2012) investigated on how player immersion and game experience are affected by this factor. They divided the challenge dimension into physical and cognitive challenge. The aim of the field experiment was to show which type of challenge has a greater impact on the experienced immersion. Additionally, the authors included the expertise of players to explore the increase and decrease of user experience if the level of challenge does not match the player's skills. The questionnaire by Jennett et al. (2008) was used to measure the degree of immersion in three field experiments. The results have shown that an increase of physical challenge does not enhance the user experience at all. In contrast, cognitive challenge increases the user experience. Furthermore, the authors conclude that a good balance between expertise of players and level of challenge is also responsible for a higher immersion. Another factor to increase immersion can be music, which is described and evaluated in Sanders and Cairns (2010). The authors found out, that music can alter the sense of time a player feels. Forty-one participants were split up in different groups playing the

same game. One part had to evaluate the game with predefined music, the others without. After the game session they had to guess their playtime and fill in the immersion questionnaire by Jennett et al. (2008). The results were not expected by the researchers. Instead of increasing the immersion rate and reducing the estimated play time which would represent a loss of sense of time, participants were less immersed than players who played the game without music. Participants who evaluated the questionnaire with a lower degree of immersion disliked the music. The researchers therefore undertook a further field experiment where participants rated and chose their favorite music before the game session started. The results indicated a higher degree of immersion with music than without. From the gained results, music can influence the degree of immersion but strongly depends on the players' music preferences. Since *Geo Heroes* is played outside independently from the player's location, fast assessment of participants can help to measure user experience unaltered. Moser et al. (2012) specialized on rapid assessment in order to measure game experience in public settings instead of just using a lab setting. The difference is that they recruited participants by placing consoles with games in a shopping mall to measure game experience after a voluntary session of playing. Due to the frequency of people in a mall, many children were interested in playing the game for a short amount of time. Through this channel, it was possible to easily recruit more than 400 children at the age of ten to fourteen as participants in a three-day event to measure game experience. Since the researchers had prior experience in this event and also in the target group, which was mainly children, they developed a questionnaire for measuring user experience in games which fits these special requirements. Game sessions only last a few minutes and children are limited in time when filling out a questionnaire. Therefore, the *Extended Short Feedback Questionnaire (eSFQ)* was developed, which is short and simple to understand, especially for children to receive appropriate feedback for measuring game experience. Elements such as fun and enjoyment but also curiosity and co-experience can be measured in a small amount of time, with minimum effort in the recruiting process of participants by using the *eSFQ*. Even if questionnaires are the most common application for evaluating user experience, there are still problems that could bias the results. Frommel et al. (2015) pointed out issues such as retrospective accuracy, the trustworthiness of users and a guessing behavior while answering questionnaires. Furthermore, the authors argue that pre-

senting a questionnaire during or after the field trial can lead to a decrease of immersion and presence in the game environment. Since these factors are relevant for measuring user experience, the researchers are tackling this problem with integrated questionnaires in games as game elements. For evaluation purposes they designed and implemented a racing game with multiple lanes. The integrated questionnaire consisted of evaluation items on each lane and participants had to change lanes in order to select the desired choice. In contrast, a control group of participants were given an overlay questionnaire, which could only be answered when switching the input controls and interrupting the game. 61 participants took part in this experiment. The results have shown that participants using the integrated questionnaire reported a significantly higher experience of presence than participants using the overlay questionnaire. The authors conclude that maintaining presence during game sessions while gathering data for field experiments is an important task to measure user experience. It is necessary to keep the players immersed. Therefore, an interruption which forces the player to leave the game world in order to fill in questionnaires should be avoided if possible. Questionnaires should only be used if other evaluation methods such as automatic detection are not applicable.

2.3 Pervasive Games

Pervasive is a broad term for describing games which connect and extend the virtual with the physical world (Benford et al., 2005a). Since location based games are using the real position of players in the game, they belong to the group of pervasive games. In literature the term *pervasive* is more often used than *location-based*, even if there is a location-based factor included. Furthermore, the design and implementation of both blur into each other and therefore this section describes pervasive games in general and their evaluation techniques.
Benford et al. (2005a) initially described pervasive games and their conjunction to traditional games. Game experience changes depending on the current context of the player. Pervasive games are therefore using (mobile) sensors to gather such information, for example the location of the user.

These type of games enable players to leave their four walls and provide an extended game experience including the real world. They also give some examples on already implemented pervasive games and describe various types. The first type is to redesign traditional computer games in a way that it will provide the pervasive factor. As an example they mention the game *Human Pacman* (Cheok et al., 2004), where players must physically run to move their character in the virtual world. The second approach is to focus on social interactions in pervasive gaming. This means that players are playing together with other players, or even non-players, in a session to reach the goal of the game. In the third type, pervasive games can help to encourage users to learn. People's motivation for learning can be unleashed when being highly physically involved in a role playing scenario in their current environment. As an example, the game *Savannah* helps children to learn about the ecology of the African savanna (Benford et al., 2005b). The last type of pervasive games are games which include commercial offerings. Proposals can be adjusted to the current context of the user, resulting in a much higher success rate. Benford et al. (2005a) analyzed these types and points out three core technologies for pervasive games. First, all of them are displaying the digital content with the help of different technologies, such as mobile phones or earphones. Second, they take advantage of wireless connectivity to communicate with servers or other players. And third, they use sensing technologies to gather information about the player's current context such as the GPS position. Benford et al. (2005a) also describes the new challenges which arise due to the nature of these kind of game play. As an example, *Dealing with uncertainty* is one of the main issues concerning pervasive games. These uncertainties are often associated with the wireless communication and mobile sensors. *Hefting domains* is also a challenge for game designers. This means that game elements consist of virtual and physical objects and therefore should be considered when they are inserted in the game design. Another problem can be *Configuration*. This issue occurs if the game should be played at many different locations and the game content has to be adapted by integrating new local information. For Benford et al. (2005a), pervasive games are exciting since there are many research challenges, but also because of their potential in commercial usage.
In the group of pervasive games, the virtual environment (VE) games are a well known representative. For example, Witmer and Singer (1998) focused on measuring presence in VE. Necessary conditions for experiencing pres-

ence are immersion and involvement. The authors describe involvement "...as a psychological state experienced as a consequence of focusing one's energy and attention on a coherent set of stimuli or meaningfully related activities and events." Measuring presence in their empirical study was done by addressing factors which related to both conditions, immersion and involvement. Based on the author's conceptual work, these factors were divided into four major categories: control factors, sensory factors, distraction factors and realism factors.
The more control a user has in the VE, the higher the experience of presence. Therefore, the degree of control is a considerable factor for the category regarding control factors. The immediacy factor is also part of the control factors and relates to the time delay between a user interaction and the consequences in the environment. The shorter the delay, the better the presence experience. Predicting what will happen next belongs to the anticipation factor which is also part of the control category. If the interaction in the VE is natural and not artificial, presence may be enhanced. Also a modification of physical objects in the VE may increase the experience of presence.
Sensory factors belong to the user's initiation of visual information but also over other sensory channels. Environmental richness can help enhancing the sensory factor by stimulating the user's visual senses. A combination of the user's senses including the consistency of these multimodal information may also increase the value of presence. The experience of presence can also be enhanced by providing users with the ability to move and look around in a VE. Distraction should be diminished in a VE to create a rich experience of presence. The user should be isolated from the physical environment by providing head-mounted displays or headphones. Also the ability and willingness of a user is necessary to reduce distractions in the physical world and therefore to receive a high value of presence. Another important factor is the awareness of interfaces. Interfaces or devices which do not suit the VE may have a negative impact on the experience of presence. Realism factors are also important in VE. These factors include the scene realism with content, textures, lights but also the consistency of information within the VE. Also, if a situation becomes meaningful to a user, the experience of presence may be enhanced. Disorientation and anxiety are also located in the realism factors. The more disoriented a user is after leaving the VE, the higher the value of presence was.
The factors are used in the two questionnaires *Presence Questionnaire (PQ)*

and *Immersive Tension Questionnaire (ITQ)*. While the ITQ is used to measure a user's tendency to be immersed, the PQ measures the experienced presence in a VE based on the factors mentioned above. However, measuring performance in VE through the manipulation of factors to experience presence is contended by the authors, but they also mention that an evidence to create a direct link between performance and presence is currently missing. McCall et al. (2004) used the ITQ including another questionnaire, think-aloud video, structured interviews and repertory grids to evaluate the presence in two VE demo applications. They conclude that the think-aloud test as a capturing method may distract participants from their actions and that post-game capturing methods such as questionnaires, interviews and repertory grids could have problems in detecting the actual presence value during the participant's sessions.

As mobile and sensor technology advances, new pervasive games have arisen in the field of research. Bell et al. (2006), for example, introduced the game *Feeding Yoshi*, a location-based game which uses public and private wireless networks as the main technology for game play. These networks have to be explored in order to feed Yoshis with fruits from plantations. The conducted field trial was split up in teams with a size of four players. Teams spread out to different urban areas and their goal was to collect more points than the other teams over a period of one week. Points were gained by feeding Yoshis with their desired food. Yoshis are non-player characters (NPCs) which can be found all over the city. The design of *Feeding Yoshi* did not force players to continuous engagement over the duration of the field trial. Instead, the game was designed in a way that it could be integrated in the daily routines of a player. During the field trial, players were observed by collecting information from interviews, game play videos, game diaries and system logs. In the results, the dependency of wireless infrastructure in cities was mentioned. The integration of the game in players' daily routines was another important factor for a successful game play. Two modes of play were recognized during the evaluation. On the one hand, players especially played the game in their free time, for example during lunch or after work. On the other hand, they integrated the game in their daily routines. The teams which were most successful managed to deal with both modes. Players also learned about their urban environment in the context of the game play and about characteristics of wireless networks. The authors suggest for a game in a long-term scenario, that there should be a connection between

game play and players' lives. Benford et al. (2006a) introduced the location-based game *Uncle Roy All Around You (URAY)*. In this game, players have to find a person called "Uncle Roy" in a public setting with a given PDA. A series of textual hints, which are often ambiguous, support the players to find "Uncle Roy" in his office. The game is fully orchestrated and additional actors are part of the game. The authors describe how such a game can blur the boundary between fiction and the everyday physical world in a mobile experience, but also point out the importance of orchestration in such a game to prevent actual and perceived risks which could harm game play. A game with less orchestration but more team oriented game play is *Can You See Me Now* described in Benford et al. (2006b). It is a location-based game with the aim that people located in a city, called runners, have to try to chase virtual avatars of players, called online-players, who control a character with the keyboard on a personal computer. The runners are equipped with a handheld computer, a WiFi antenna, a GPS module and a walkie-talkie. On the handheld device a map of the city with their own and other players' current positions is displayed. The online-players use their personal computer connected to the game session via internet to also display their and other players' current positions on a map. Online-players can chat with other team members about tactics. One interesting feature in the game mechanics is that runners can see the chat of online-players on their handheld device and online-players can hear the talk of the runners. The authors focused on strategies to deal with uncertainties and divided them into five categories: Remove, hide, manage, reveal and exploit. They also tackled some technical issues by applying some simple improvements, for example a heat map to display the strength of GPS signal for runners. O'Hara (2008) presents findings from a long-term (three weeks) field study of *Geocaching*. Players who had already played the game voluntarily were chosen as participants. The authors focused on the location-based experience not just in consuming, but also in creating this experience. Another point of their work is the combination of physical and virtual elements in game play. In their field experiment, participants had to fill in a questionnaire and had to take part in some interviews. One interview established background context about participants and was applied before the trial took place. Another interview included questions about *Geocaching* experience and the underlying motivations. The results differ from the experience factor and characterization of players. Inexperienced cachers' behaviours were just based on finding a

particular cache or to log a find, which was very infrequent. The enthusiastic cachers integrated their *Geocaching* activities in their daily routines. These type of players also monitored these activities through different channels (e.g. website, email notifications), which was important for keeping up motivation. The authors distinguish between different factors of motivation. One of these factors is the social walk, where the primary motivation is the possibility to go out with family or friends. The GPS device is only used for navigation purposes and nature can be enjoyed instead. Another type of motivation is to discover and explore places in the home town or in a new environment. Collecting is also a major incentive in this location-based game. The collection of caches represents the success of a player and is a strong motivational factor for ongoing participation. But there are also other reasons for an increased motivation such as a public profile with statistics or the challenge and competition in the community. Another stimulative factor to play *Geocaching* is the creation of caches. Players either want to give something back to the community or others want to share significant places. Furthermore, players can add some historical, geographical or personal information to a cache, while others have a demand in gaining reputation of the community when creating a good cache.
Since "place" has a significant role in games such as *Geocaching*, Reid (2008) describes the impact of place and "design for coincidence" in the design of a location-based game. She also reveals fundamental characteristics of such games. When designing a location-based game, the size and duration of the game can vary and have an influence on game play. Also the infrastructure, which includes the capability of the device, the network availability and sensing technologies, has an effect on how the game should be designed. Another characteristic is the significance of place and its role. "Place" has been described in Harrison and Dourish (1996) as a location with persistent social meaning, and Dourish (2006) describes how the mapping from virtual worlds to the physical presence in the real world can extend the access and interaction of such places. Reid (2008) classifies location-based games in relation to the way their virtual worlds interfere with the real world. The first class covers the selection of the game area by the user, whereas the second class uses the game's infrastructure. In the third class the game designers choose the location, since specific properties are often required in the real world for game play. Design by coincidence is also a relevant part, according to Reid (2008):

> *"Coincidence is defined as being the noteworthy alignment of two or more events or circumstances without obvious causal connection. When a player experiences such a coincidence it feels 'magical' and thus leads to feelings of wonder and excitement."*

Different forms of coincidence exist, such as natural, social and feigned coincidence. Those are examined by the author and how the knowledge of the environment can be used in order to increase the chance of creating such moments is described.
A more physical activity based game compared to others is described in Soute et al. (2009). They compared two versions of the same outdoor pervasive game in a case study. The game is called *Save the safe* and is categorized as a *Head-Up Game*. Players are split up into two teams: burglars and guards. The goal is to protect the key of a safe from the burglars. The game ends when either the guards have successfully protected the key from the burglars for five minutes or the burglars stole the key and unlocked the safe. There were two versions for game play. The first version was played with a belt which vibrated when the wearer obtained the virtual key. Passing the key was done by coming close to the key's owner. The belt also displayed the team color. In the second version, a physical ball was used as the key and children could throw the ball to other team members or the opponent team could intercept to obtain the key. In this case the belt was also used but only for indicating the team color. 27 children between eight and nine years were chosen to play the game. The goal of the study was to evaluate the social interaction and the physical activity of the participants during the field trial. During the game sessions the children were observed with video cameras. Afterwards, participants were taken inside to fill in a questionnaire and to rank elements of the game through *Likert* scale options ('least fun' to 'most fun') in a group discussion. The authors conclude from the gained results that participants preferred the virtual game. Since the vibrating bell was a newly introduced technology, the results could be biased and one should not conclude prematurely the case for virtual objects. Children also mentioned that the physical interaction with the physical object was easier. Another point which they conclude is that *Head-Up Games* can support physical as well as virtual game objects but relative advantages should be explored in the design process.
Since immersion is a relevant factor in games, Waern et al. (2009) discuss

the creation of a fully immersive game by analyzing *Nordic Style Live Action Role-Playing (Nordic LARP)* to find out what is important to create a *Three-Sixty Illusion* for users in order to apply such techniques to pervasive games. Nordic LARPs do not use any technical devices. Participants are trained to play their character's role to perfection. The real world is physically transformed to the content of game play, which results in an authentic environment. Furthermore, activities are not just simulated in LARPs. Instead they are performed for real, which leads to a physically and emotionally immersive role-playing. The authors describe the *Three-Sixty Illusion* in *Alternate Reality Games (ARGs)*, a subclass of pervasive games, which are suitable for mass-participation and playing without environmental restrictions. The players of ARGs are often confronted with a narrative story they are able to uncover during game play. The authors conclude that pervasive games are played in the ordinary world and should concentrate on the creation of an illusion in the real world. A pervasive game should therefore map the virtual game content on a physical landscape. To prove their arguments they used two games called *Momentum* and *Interference*, which were designed to create a *Three-Sixty Illusion* for participants in a field trial. Since technical devices were used in both games, the authors also conclude that those should be created to fit to the game and possible failures should be actively involved in the game.
Another location-based game with focus on the relation between virtual and physical world is described in Carrigy et al. (2010). They designed and evaluated a location-based mobile game (LBMG) called *Viking Ghost Hunt (VGH)*. The story of the game covers a Gothic ghost story set in Dublin. The players slip into the role of a paranormal investigator and try to solve the myth by hunting ghosts and collecting hints. The game is story driven and players have to unlock narrative parts in order to progress. In the design of the game, the authors concentrated on the aesthetics of role-play and on the mapping between the real and the virtual world to create an immersive and engaging experience. Augmented reality (AR) was also used to create a deeper immersion during game play. The game was evaluated by 19 participants over a series of field trials in a three-day period. Feedback on engagement was obtained through a questionnaire including bounded and open-ended questions. During game play data was logged through system logs. The evaluation exposed that narrative and imaginative immersion increases when players gradually understand the game mechanics. The

authors conclude, that the selection of the location has also an impact on the engagement and immersion. Hence, the location should have a relation to the content of the game world. Furthermore, a technology such as AR can also enhance the gaming experience through the creation of a hybrid game world, and thus providing a deeper immersion to the game.
Lund et al. (2010) created a location-based mobile game based on the concept of *Geocaching* including exploration and discovery as core mechanics as well as a competitive game play. They describe that in *Geocaching* there is lack of competitive elements and therefore it is categorized rather in the sectors of activities than in games. For this reason they designed and evaluated a game called *Big Game Huntr*, where users can create events around a chosen topic in a defined location. Events are specific challenges, where players have to take a photo of their solution and upload it to the community for approving that the challenge has been solved. Points are given from the creator of the event by rating the solutions. The goal is to get as many points as possible by providing good solutions. At the end of the game, a leader-board displays the sum of each player's points to select a winner. The game was played by 53 users in a time frame of five hours, who created 56 events and uploaded 378 photos. The authors show with these numbers how fast user-generated content can be created in a short amount of time. They logged and plotted the movement patterns of participants and conclude from these results that prior knowledge of the area where the game takes place leads to a more effective game play.
In contrast to previously mentioned works, Tiensyrjä et al. (2010) designed a light-weight location-based multiplayer game, which was implemented as a web service. It therefore does not need an installed application, just a web browser on the device and a Wireless-LAN connection fulfills the requirements to play the game instead. The goal of the game named *panOULU Conqueror* is to conquer wireless access points in a city to score points for the team. The game was evaluated through a tournament with a duration of four weeks including 96 players in 31 teams. The authors conclude that the technical design as a web service is an advantage, since no specific soft- or hardware is needed. Also, players managed to deal with the uncertainty of the positioning mechanism due to the visual feedback implemented as a notification at the game client. As mentioned in their later work, the authors wanted to create a higher level of pervasiveness through game elements which can be found in the real world.

Despite the progress over the years in the development of pervasive games, most of them require a display. Kurczak et al. (2011) claim that location-based games have two major problems in this area. First, a player is required to monitor the display of the screen most of the time and accidents may occur. Second, using screens during game play leads to a disconnection of the physical and virtual world. Therefore, they designed a location-based game, but in contrast to others they use *ambient audio* to navigate the player through the physical world. The game *Growl Patrol* involves escaped small animals from a pet shop and a tiger. The escaped animals move around and have to be caught by the player before the hungry tiger finds them. Stereo headphones are used to hear the noises of animals. Depending on the direction and position of the player, a continuous audio stream sounds left or right to determine the animals' positions. The more animals the players rescue without losing any to the tiger, the more points they get. The goal of the field study was to test whether the usage of ambient audio increases the immersion and safety in location-based games. 24 participants took part in the field trial with an average age of 25 years. The authors compared an ambient audio interface with a visual interface. They used the questionnaire by Witmer and Singer (1998) and a semi-structured interview for evaluation. They conclude from the results that players who used the version with the visual interface achieved higher scores, but also had more collisions with the environment than in the ambient audio version. The presence value in the ambient audio version was also significantly higher and participants reported that the game was more fun when playing with the audio interface. The authors reveal some limitations of the usage of ambient audio in location-based games such as the inaccuracy of mobile sensors and the number of audio streams a player can track.
In order to test sounds, texts and graphics in pervasive games, Díaz et al. (2014) developed and evaluated the game *Mystery of Elin*, a game played by children on tablets to learn about cultural heritage. The game includes a story which can be progressed by solving various tasks. The goal of the evaluation was to measure the game experience of participants in a field trial. The authors used quantitative and qualitative methods to evaluate the game. The amusement was measured with a four level scale from 'boring' to 'very exciting'. The game experience was also evaluated during interviews and plotted in a four level scale from 'not funny' to 'exciting'. The authors conclude that there are limitations in a location-based story driven game.

Due to the fact that these games are played outside, sunlight can have a direct influence on the device's screen. Also sounds can be missed in an outdoor environment. Therefore, graphics can be more important and should be designed in a way that they can also be noticed by players under sunny conditions.
Since players have to overcome a certain distance in location-based games, Procyk and Neustaedter (2014) evaluated user experience for this characteristic of a location-based game called *Geolocated Embedded Memory System (GEMS)*. It was designed to create and share narrative entries to friends and family members. Players can visit and tag meaningful places with text and photo for others. Evaluating *GEMS* was done in a user study over a period of three weeks with 15 participants. Pre-play, follow-up interviews and diaries were transcribed to gather qualitative information. The results revealed that the number of players willing to create and collect entries was much lower than the authors expected. In the interviews participants explained that visiting places which are faraway were harder to get to in daily life. The required time had a major impact on creating or consuming stories at distant locations and was a discouraging obstacle during game play. The authors suggest for location-based games to support alternative methods for creating and consuming content in addition to reduce the immense effort in order to keep the player's motivation high. They also conclude that intrinsic as well as extrinsic motivation are necessary to guarantee the creation and consuming of game content.
Due to the fact that *Geo Heroes* uses map information to place content in the virtual world, an interesting paper is presented by Spiesberger et al. (2015). The authors designed and evaluated a game called *Woody*. *Woody* is a pervasive game which uses an open data collection of trees in cities for creating game content. The goal of the game is to interact with virtual trees, which are placed at the position of real trees in the current environment, using various physical transformation gestures with smartphones to keep a timber worm alive. User experience and outdoor activity was measured during the evaluation. The authors therefore divided the 38 participants at the age of 11 to 12 years into two groups. Each group received a different version of the game. One group should evaluate the location-based version of *Woody*. The other group received an indoor version of the game. Quantitative information was collected with *Google Analytics* during the 12 days of the field trial. The results showed that the level of activity and the

time spent outdoors was much higher for the location-based variant. The received qualitative feedback was very positive for participants with the outdoor version. They voluntarily altered their daily routines to progress in game play. The authors are planning to evaluate the game again in larger field trials with focus on user experience, game strategy and long-term behavioral effects.

2.4 Designing Location-Based Games

Since this thesis describes the design and implementation of a location-based game, some literature was reviewed describing the most important factors for creating such games. For example, Bowser et al. (2013) introduced *PLACE*, an approach for location-based applications and games. It supports the usage of different levels of **P**rototyping in the sections **L**ocation, **A**ctivities, **C**ollective experience and **E**xperience over time. The method provides an iterative design process by using a set of key principles. First, start with a low-fidelity prototype and reuse its content to build a high-fidelity prototype. Second, designers should observe participants and treat them as co-designers during a field trial and integrate the provided feedback into the application. The third principle describes the testing with locations. Even if the application is in an early stage, the prototype should be tested at a location, even if the size differs from the finished application's location. In the fourth principle, designers are prompted to lay the focus on activities, user motivation and game dynamics instead of the design of interfaces. The social experience is the fifth principle of PLACE. If the application supports social interactions, the prototypes should also include them to create a more authentic experience in this area. The last principle should consider time authentically. The authors argue that even if prototypes are time constrained, they should reflect the duration of the finished applications as precisely as possible. A game which supports spontaneous play and planned expeditions should consider the evaluation for both time periods. With the mixed-fidelity prototypes and the applied principles, designers gain a better understanding of a user's requirements during the development process, resulting in an improved user experience in the final application.

Neustaedter et al. (2013) describe a design guidance for scaleable location-based games in form of lessons learned by studying the well-established game *Geocaching*. Over two years, participation and observation was studied through an online survey with many participants. As a result of the study, the authors present seven lessons which can be applied to location-based games. *Lessons 1 and 2* describe the lightweight and more elaborate creation of game elements for players. In *Lesson 3*, players should learn the game elements and content over time to obtain new types of experience. It should be easy to report and monitor game elements through activities, which is described in *Lesson 4*. *Lessons 5 and 6* describe the observation and interaction with other players and non-players, which should be possible through online activities. In *Lesson 7* the possibilities are presented players should have to maintain and remove game elements which were created by other players. The authors also introduce some characteristics which can vary in location-based games. Depending on the game, the content can be fixed or continual. The interaction with game elements can be easy or complex and the creation can be done through an admin or player, or even by an algorithm. A further classification is the interaction between players, which can be performed synchronously or asynchronously. The authors mention that it is easier to apply the provided lessons after identifying the characteristics of a location-based game.

3 Research

In this thesis a location-based game was designed and implemented for *Android* devices. The game is named *Geo Heroes* and narrates the journey of a medieval knight in a fantasy genre. The game is a typical role-playing game played on a mobile device combined with location-aware technologies, such as GPS, compass and the real world map as the virtual world. A rudimentary prototype of this game was initially developed in a previous work by the author of this thesis.

3.1 Research Question

The aim of this thesis is to improve *Geo Heroes*, which uses map information to place game content on the map in an user-centered design approach. As described in the related work above, there are many approaches to measure user experience or player enjoyment for traditional or pervasive games. The question is, which technique fits *Geo Heroes* best and how applicable the chosen methods are to measure user experience. Additionally, the results of these approaches should also point out concrete problems to provide tasks for the next iteration in design and development. Furthermore, this game was designed to be played by experienced gamers of traditional computer games, but also for people who never or rarely play games. Besides, analyses and conclusions in developing and designing location-based games from the related work section may be relevant in order to prove the gained results.

3.2 Hypotheses

In order to answer the formulated research question, a list of hypotheses is used to subdivide integrated topics.

H1: Convenient quantitative methods, such as questionnaires, provide a good estimation to measure user experience for *Geo Heroes*.

H2: Convenient qualitative methods, such as feelings or interviews, indicate whether players of *Geo Heroes* like or dislike the game in specific situations or at all.

H3: Specially designed quantitative methods for location-based or pervasive games are more suitable than quantitative methods for traditional computer games.

H4: It is easier to provide evidence of a lack in user experience when results of qualitative methods are used to support the results of quantitative methods.

H5: An algorithm which uses map information to place virtual game content at appropriate places in the real world increases the overall user experience.

H6: Players who are familiar with playing traditional games will have easier access to the game and hence a better user experience.

3.3 Specific Aims

The specific aims are used to test formulated hypotheses. Each specific aim is matched to one or more hypotheses:

An empirical study is conducted which is evaluated by pervasive and traditional quantitative and qualitative methods in order to gain results in measuring user experience (H1,H2). Additionally, the applied methods will be compared and discussed to see whether they aid one another (H4) and which are the most valuable for a game such as *Geo Heroes* (H3). The empirical study provides a field experiment with participants with and

without gaming experience (H6). Furthermore, the participants will be split up in an experimental and a control group. The experimental group will play the game with the algorithm based on map information, whereas the control group will play *Geo Heroes* with a simple algorithm based on a random principle to place virtual content in the physical world (H5).

4 The Game

Geo Heroes is the second version of the existing prototype *Heroes on Earth*, which was already developed by the author of this thesis. *Geo Heroes* is a location-based mobile role-playing game (LBMRPG) in a medieval-fantasy setting. At the moment the game can only be played on Android devices which support GPS and compass. An iOS or Windows phone version was not developed since it would exceed the time constraints of such a thesis. In contrast to its 3D predecessor *Heroes on Earth*, *Geo Heroes* is now implemented as a pure 2D game to overcome performance and memory issues. Also multithreading is now supported for generating and placing content. Figure 4.1 shows an in-game screenshot of the current implementation of *Geo Heroes*.

4.1 Game Design

As previously mentioned, the game is a location-based game which is played on mobile devices. Users have to turn on the GPS signal of their devices to locate their current geographical position. The player's avatar is represented by a medieval knight with the goal to become stronger by defeating enemies or collecting resources to level up. The world the knight lives in is the current geographic map of the player. The map styling is achieved by using the map and the editor of *MapBox*[1]. The GPS position of the device is used to control the movement of the knight on the map. In the environment, enemies, resources and shops spawn around a certain area of the player's avatar. If a player is close to such an object, a button is shown. Clicking on the button enables an interaction between player and object.

[1] http://www.mapbox.com/

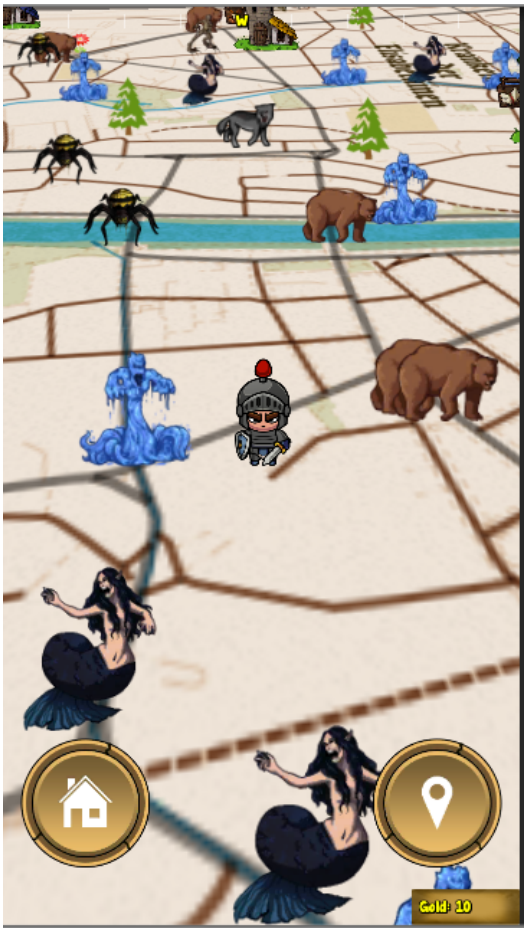


Figure 4.1: An in-game screenshot of *Geo Heroes* captured in the city *Graz*.

Interacting with an enemy starts a turn-based battle. The player should use the provided skills and attacks to defeat the enemy to receive experience which is necessary for the game's progress. There are two outcomes after a battle has started: Either *victory*, which is achieved by reducing the enemy's health points to zero, or *defeat*, which is encountered when the player's health points fall to zero before the enemy's one. Enemies are divided into five categories: nature, grass, mountain, spiritual and normal enemies. Each enemy differs according to its current value of status attributes such as health, attack or defense values. The higher these values, the stronger the combatant. Status values increase when the enemy or player has a higher level. This is achieved due to the different implemented status development of each type of combatant. An overview of all status values with its meaning is shown in Table 4.1.

A player can also interact with resources. There are three types of resources: ore, herb and wood. Collecting resources is implemented with a random function, which represents the luck of a player. When a resource on the map is collected, the player receives zero to three pieces of that resource depending on the outcome of the random function. Resources are required to craft equipment and to brew potions. Equipment can also be gained from the loot of a winning battle, but the probability in this case is low. The player is therefore supported with a crafting system to create appropriate equipment which increases the knight's status values when equipped. For this thesis, an alchemy and forgery system was developed. Alchemy is used to brew potions with herbs to refill consumable status values during a battle such as health points. Forgery is the ability to create equipment such as armor and weapons out of ores and wood to increase static and consumable status values.
Since an in-game currency exists in each role-playing game, *Gold* was introduced in *Geo Heroes*. *Gold* can be earned by winning battles and can be redeemed in *Shops*. The currency can then be used to buy items or equipment for the player's inventory in four different types of shops. *Doctors* are selling medical aids like potions, which regenerate *Health* or *Mana Points*. *Equipment Stores* have armors and weapons to increase the status values of players. *Taverns* are selling food items, which restore consumable status values such as *Health Points* outside of battles. In *Educational Institutes* players can buy recipes to craft new items. A full list of game content with their

Status Attribute	Type	Description
Health Points (HP)	consumable	Represents the life of a combatant. Reducing the life points to zero means the death of the combatant.
Mana Points (MP)	consumable	The attribute is used to cast spells.
Physical Attack (ATK)	static	The physical attack increases the physical damage of a combatant's attack.
Magical Attack (MATK)	static	The magical attack increases the magical damage of a combatant's attack.
Physical Defense (DEF)	static	The physical defense reduces the incoming physical damage of an enemy.
Magical Defense (MDEF)	static	The magical defense reduces the incoming magical damage of an enemy.
Mana Regeneration (MREG)	consumable	Regenerating Mana over time is connected to this attribute and necessary to recover the consumable Mana Points.
Speed	static	Speed is used to determine the order a combatant is allowed to make an action in the turn-based battle.

Table 4.1: The table presents the status attributes in Geo Heroes, which are part of each combatant, with a short description of its use.

appropriate type can be viewed in Table 4.2.

4.2 Implementation

Geo Heroes was implemented with the game engine *Unity*[2] and the help of various frameworks, plugins and third-party resources to accomplish the completion of a playable prototype in a comparatively short amount of time. Since the game is played on mobile devices, the usage of different sensors had to be implemented additionally. The following paragraphs are introduced to describe each important implementation aspect in detail.

Framework As mentioned before, *Geo Heroes* uses frameworks for a faster development. *Unity* as its game engine supports plugins, called *assets*, which are available in their own asset store[3]. *Unity* was chosen due to its easy deployment process for mobile devices. For this game, two major assets were used: *Online Maps*[4] and *ORK Framework*[5]. *Online Maps* supports the rendering of map tiles, GPS handling, marker support and many other features necessary for a location service. *ORK Framework* is a highly configurable role-playing game (RPG) engine inside *Unity* which supports all the RPG functionality required for this project. Both assets were combined to create the base for a location-based mobile RPG. Other assets were also used such as images, sounds and music of professionals to provide fundamental aspects of a game which were partly used for answering the research question.

Objects and Markers *Geo Heroes* uses the GPS position to display the player's avatar which is centered on the map view. Objects, such as enemies, resources and shops are also placed on the map. This can only be achieved when each object has its own geographical position. Assigning geographical

[2]http://www.unity3d.com
[3]https://www.assetstore.unity3d.com
[4]http://infinity-code.com/en/products/online-maps
[5]http://orkframework.com/

Game Content	Type	Name
Enemies	Spiritual	Skeleton
		Ghoul
	Normal	Minotaur
		Mantikor
		Ogre
		Harpy
	Vegetation	Wolf
		Bear
		Giant Spider
		Panther
	Mountain	Griffin
		Golem
		Fire Dragon
		Bat
	Water	Siren
		Water Spirit
		Giant Crab
Resources	Orb	Copper Ore
	Herb	Camille
		Violet
	Wood	Pine Wood
Shops	Marketplace	Marketplace
	Equipment Store	Equipment Store
	Education Store	Education Store
	Tavern	Tavern
	Doctor	Doctor

Table 4.2: The table presents the game content, such as enemies, resources and shops. Game content can be divided into content types. A concrete game object such as a *Skeleton* belongs to a type.

coordinates to each object is therefore a relevant implementation detail to calculate distances for the interaction radius to other game objects.

Geofences and Local Notifications The problem with *Unity* applications on mobile phones is, that the application freezes when it is moved in the background or the screen is turned off. When playing a location-based game, it is often not necessary to always display the screen, since players have to walk for a while until a certain event or interaction is in close proximity. Hence, local notifications were implemented with a native library developed with *Java* for *Android* smartphones. The purpose of the library is to start a background service which is responsible for notifying the user when an object is in their surrounding until the game is actively running again. With the usage of *Geofences*[6] local notifications can be triggered when entering or leaving such an area. Hence, such a notification is shown when entering a *Geofence* on the normal user interface outside of the application and destroyed when leaving the *Geofence*.

Procedural Content Generation and Map Information Location-based games are often constrained in their locality or the provided content is predefined by game designers to reduce the complexity of the game's logic. *Geo Heroes* does not restrict the player to a specific location with static content. Instead, it uses procedural content generation (PCG) to provide an ongoing experience all over the world. Since spawning content over the area of the whole world at the same time does not make sense in a single-player game, a dynamic spawning algorithm was designed and implemented. Moreover, the game uses map information provided by *OpenStreetMap (OSM)*[7] to place game content to an appropriate type of location.
Three data representation types of *OSM* are used in *Geo Heroes* to position game content: *Nodes*, *Ways* and *Areas*. A *Node* represents a single point and contains a node id and the longitude and latitude coordinates. Furthermore, a node has tags to define its purpose. Tags are defined with a key and a value. For example, a tree may be tagged with **natural=tree**. A *Way* is defined by a sequence of nodes, such as a highway or a river. In contrast

[6] https://developers.google.com/android/reference/com/google/android/gms/location/Geofence
[7] http://www.openstreetmap.org

to a *Way*, an *Area* represents a closed way, where start and end point are identical. An example for an area is a park, which would be tagged with **leisure=park**.
In *Geo Heroes* the map information is used to map different types of enemies, resources and shops to an appropriate *OSM Node*. A full map of relations between game content and *OSM Tags* is provided in Table 4.3 for enemies, Table 4.4 for resources and Table 4.5 for shops.

Enemy Type	OSM key	OSM value
Water	natural	water, wetland, bay, beach, coastline, spring
	waterway	riverbank, stream, canal, waterfall, dam, dock, weir, beach
	leisure	beach, fishing
Vegetation	landuse	forest, grass, meadow
	natural	wood, scrub, heath, grassland, fell, bare_rock, scree, sand, mud
	leisure	garden
Mountain	natural	peak, volcano, valley, saddle, ridge, cliff, rock, stone, cave_entrance
Spiritual	amenity	crypt, crematorium, grave_yard, place_of_warship
	landuse	cemetery
	historic	ruins, rune_stone, tomb
Normal	leisure	pitch, park, play_ground, adult_gaming_centre, amusement_arcade, bandstand, dance, golf_course, stadium

Table 4.3: The table shows the mapping between types of enemies in the game and their appropriate OpenStreetMap key and value pairs.

Resource Type	OSM key	OSM value
Orb	natural	peak, volcano, valley, saddle, ridge, cliff, rock, stone, cave_entrance
Herb	landuse	grass, meadow
	natural	scrub, heath, grassland,
	leisure	garden
Wood	landuse	forest
	natural	tree, tree_row

Table 4.4: The table shows the mapping between types of resources in the game and their appropriate OpenStreetMap key and value pairs.

Shop Type	OSM key	OSM value
Doctor	amenity	clinic, dentist, doctors, hospital, pharmacy
Tavern	amenity	cafe, supermarket, fast_food, restaurant, pub, bar
	shop	bakery
Marketplace	amenity	marketplace, supermarket, townhall
	leisure	park, playground
Equipment Store	shop	clothes, shoes, boutique, fabric, fashion, leather
Education Insitute	office	educational_institution
	amenity	townhall, kindergarten, library, public_book_case, school, university, college

Table 4.5: The table shows the mapping between types of shops in the game and their appropriate OpenStreetMap key and value pairs.

5 Evaluation

The goal of this thesis is to improve the gaming experience (GX) of *Geo Heroes*. The traditional way of improving the user experience is to use user-centered design. This approach is based on an iterative process which covers the formulation of requirements, the creation of prototypes and the evaluation with end-users. After each iteration, the development team implements the gained information into the game. This process aims at the continuous integration of end users in the product cycle, which ensures high quality and a good user experience.
However, measuring this value for a location-based game is more complex than for traditional games or 'productive' software. The evaluation of a software which was designed to improve productivity will focus on effectiveness, efficiency, learnability and memorability. These factors can also be applied to games, but a major factor is still missing for games, which is *fun* (IJsselsteijn et al., 2007).
Furthermore, a location-based game is played in a large outdoor area. This means that an evaluation in a user-centered design approach should be done in an equivalent setting while assessing participants' experience during game play(Bowser et al., 2013).

5.1 Evaluation Tools

The goal of an evaluation which measures user experience in games is to simply remove the obstacles to fun (IJsselsteijn et al., 2007). Therefore, methods are used which point out positive and negative experiences of a

user while playing *Geo Heroes*. In this thesis three different approaches from literature were applied to measure the end user experience in this game. Two approaches are based on questionnaires in order to evaluate the user's gaming experience. The third method is using in-game capturing to record players' emotions.

5.1.1 Game Experience Questionnaire

IJsselsteijn et al. (2007) describe the conception of a questionnaire which measures the user experience in digital games. This questionnaire is independent of gaming genres and can be adopted to any kind of digital games. This approach not only includes traditional usability metrics, it also covers immersion and flow as essential factors for evaluating game experience.
The questionnaire is modularly structured and consists of three modules:

Core Module The *Core Module* contains 33 'Likert-scale' questions from zero (not at all) to four (extremely). These questions are categorized in seven components:

- Competence
- Sensory and Imaginative Immersion
- Flow
- Tension/Annoyance
- Challenge
- Negative affect
- Positive affect
- Social engagement

These questions assess the gaming experience while playing the game and therefore should be applied immediately after the game-session has finished.

Social Presence Module The *Social Presence Module* is used to assess the user's psychological and behavioral involvement with others. Since *Geo Heroes* does not include multiplayer or social interactions, this module is skipped in the evaluation.

Post-Game Module Assessing players' feelings after they have stopped playing a game can be achieved with the *Post-Game Module*. With this module a naturalistic gaming behavior can be assessed. An example would be the decision to play the game voluntarily. The module consists of four components:

- Positive Experience
- Negative Experience
- Tiredness
- Returning to Reality

Since these questions assess the user's feelings after the game-session, this module will be conducted after the *Core Module*.

5.1.2 Pervasive Game Flow Questionnaire

Since this thesis describes the evaluation of a location-based game, a further method particularly for these type of games is used for evaluation. While Sweetser and Wyeth (2005) introduced the *Gameflow Model*, which is used for understanding player enjoyment in traditional digital games, Jegers (2007) extended this model especially for pervasive gaming (*Pervasive Gameflow Model*). Both approaches are based on the work of Csikszentmihalyi (1990), where the term *Flow* was initially introduced. Here *Flow* is defined as the optimal experience in enjoyment independent from culture and activity. According to the author's study, the optimal experience or flow consists of eight elements:

1. a task that can be completed
2. the ability to concentrate on a task
3. the concentration is possible because the task has clear goals

4. the concentration is possible because the task provides immediate feedback
5. the ability to exercise a sense of control over actions
6. a deep but effortless involvement that removes awareness of the frustration of everyday life
7. concern for self disappears, but sense of self emerges stronger afterwards
8. the sense of the duration of time is altered

Combining these elements creates such a positive experience (flow), which rewards people in a way that they do not need any additional rewards, for example money. Since the eight elements define *Flow* without the context of digital games, Sweetser and Wyeth (2005) build the *Gameflow Model* in their study upon these items. Combining comprehensive review of literature on usability and user experience in digital games with Csikszentmihalyi (1990) model of *Flow* finally resulted in the *Gameflow Model*. Sweetser and Wyeth (2005) started with the mapping of elements between game literature and the elements of flow. In Table 5.1 the transformation between these elements is shown. Since "a task that can be completed" is the game itself, it is not directly part of the evaluation in the *Gameflow* model, whereas other components are strongly coherent. Players should be confronted with tasks that have clear goals and a high work-load, while still being challenging to be enjoyable. This requires concentration on a task's process and with received feedback a player will feel control over this task, which finally results in immersion. Immersion is the state of being effortlessly involved in a task or activity while losing awareness of the real world. The social interaction component is not related to the flow criteria but Sweetser and Wyeth (2005) argue that in game literature this element is also responsible for achieving a high user experience and therefore it is mapped to the *Gameflow Model*.
Additionally, games which consist of the elements of flow should also bear a good ratio of player skills and provided challenges in the game (Schell, 2008). People like challenges, but remaining too long on the same difficulty level can cause boredom as people's skill rises while performing a task. On the other hand, if a challenge cannot be achieved, frustration is triggered and people start to look for other activities which give them this kind of feeling. Keeping the balance between frustration and boredom is called

Games Literature	Flow
The Game	A task that can be completed
Concentration	Ability to concentrate on the task
Challenge Player Skills	Perceived skills should match challenges and both must exceed a certain threshold
Control	Allowed to exercise a sense of control over actions
Clear Goals	The task has clear goals
Feedback	The task provides immediate feedback
Immersion	Deep but effortless involvement, reduced concern for self and sense of time
Social Interaction	n/a

Table 5.1: The table, adapted from Sweetser and Wyeth (2005), represents a mapping between the components of the *Game Flow Model* and the characteristics of *Flow*.

the *Flow Channel*. A visual representation of the *Flow Channel* is shown in Figure 5.1.

However, the standard *Gameflow Model* by Sweetser and Wyeth (2005) is used for traditional computer games, which are typically played in indoor locations. Pervasive and hence location-based games differ from these kind of games as they are often played outside and include objects and subjects of the real world. For this reason Jegers (2007) introduced the *Pervasive Gameflow Model* which is based on the *Gameflow Model* but considers attributes specific to pervasive games such as mobile and place-independent gameplay, social interactions between players and integration of the physical and virtual worlds. As a result, this study establishes a model for evaluating player enjoyment in pervasive games. As Jegers (2007) proposes in his conclusion and future work, the model needs to be empirically validated with user-centered evaluation of various pervasive games for further refinements and improvements.

The *Pervasive Gameflow Model* was used to evaluate player experience after playing *Geo Heroes*. Each element, for example *Concentration*, has a number of criteria defined by Sweetser and Wyeth (2005) and Jegers (2007). These

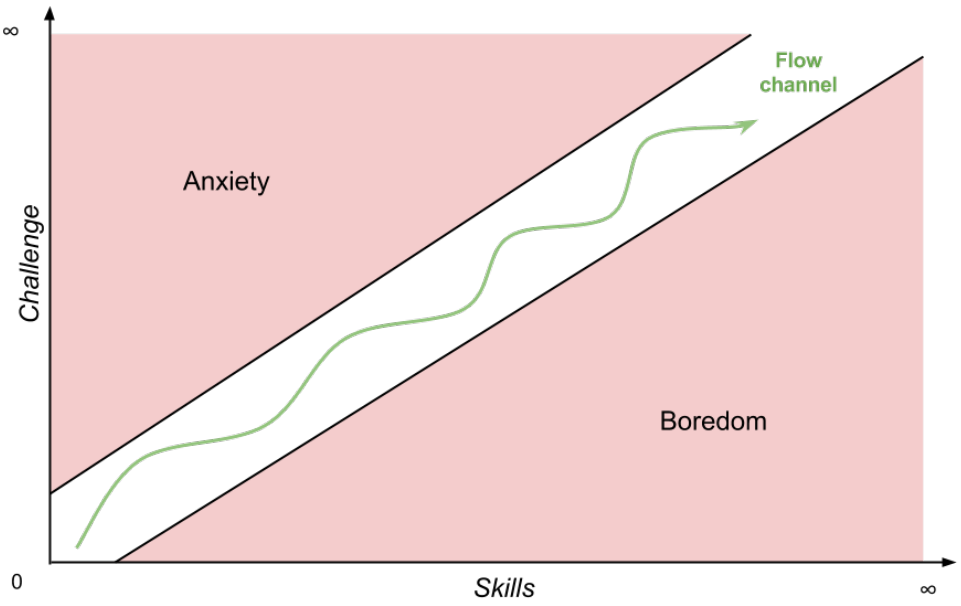


Figure 5.1: The figure describes the Flow channel adapted by Schell (2008). Flow exists, if the balance is kept between player skills and challenges in the game. When players are in this channel, they feel motivation and the chance of recurring play is high. If the challenge is much higher than their skills, players are frustrated. In contrast, when their skills are much higher than the provided challenges, players feel bored.

criteria are adapted and transformed to questions. An example criterion *"players shouldn't be burdened with tasks that don't feel important"* is transformed to *"I was just confronted with tasks that felt important"*.
All criteria of all components are reformulated in this way to generate a questionnaire out of it, except the *Social Interaction* element. This one was removed from evaluation since *Geo Heroes* does not provide multiplayer or social interactions at the moment. The remaining elements consist of 34 'Likert-scale' questions from zero (not at all) to four (extremely).

5.1.3 Emotion Wheel

Although questionnaires deliver the participant's subjective reports in a standardized way, a further approach for measuring player enjoyment during game-play session is applied for this evaluation. In contrast to a questionnaire, real-time reports for measuring gaming experience can capture the player's feelings immediately when they occur. This means that participants do not have to remember emotions they feel during game-play after a game-session and feelings can be accessed and logged directly. The approach is based on Baillie et al. (2011), where an *Emotion Wheel* is used for reporting players' feelings (Figure 5.2). This wheel consists of 24 distinguishable emotions arranged in a circle. The emotions are categorized in four quadrants:

- I: pleasure/high arousal (excitement)
- II: displeasure/high arousal (distress)
- III: displeasure/low arousal (depression)
- IV: pleasure/low arousal (relaxation)

However, the *Emotion Wheel* was applied at the end of their conducted field experiment and could be problematic since players may not remember all of their emotions.
The approach used in this thesis will tackle this problem and integrate the demand of players' emotions directly into the application after each interaction. Players are then forced to report their feelings immediately after interacting with objects in the game world.

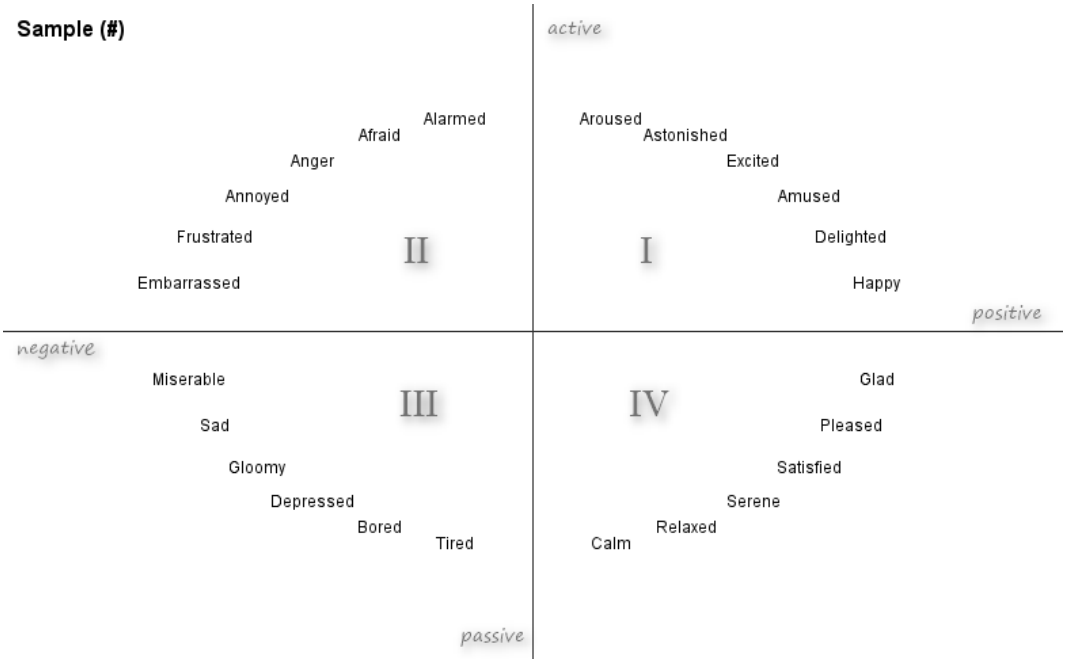


Figure 5.2: The figure shows the base concept of the emotion wheel. 24 emotions are listed, whereas the horizontal axis splits the active and passive emotions, and the vertical axis the negative and positive ones. Roman numerals divide the wheel into four quadrants: I...active-positive, II...active-negative, III...passive-negative, IV...passive-positive. The sample text will be replaced with the outcome of an object interaction and the hash-symbol with the number of samples. Figure is adapted from Baillie et al. (2011).

5.2 Participants

The registration for applying to the field trial was done through social media posts (FB, Twitter) and newsgroup (news.tugraz.at). In the end, ten people had registered for the experiment. From these ten participants, nine successfully took part in the field trial (one of them got sick), where seven were male and two female. The average age was about 30 years with a standard deviation of 3.75. All of them applied voluntarily, but as an expense allowance three *Amazon* vouchers, each worth 15 Euros, were raffled after the field trial.

5.3 Evaluation Procedure

In the previous section the methods for evaluating user experience in *Geo Heroes* were described. This section focuses on the direct application of these methods. Since *Geo Heroes* was designed with an user-centered design approach, the evaluation consists of a field experiment with participants playing the game in a defined session. The field trial took place in the city centre of *Graz / Austria*. The city has a high number of *OpenStreetMap* nodes and ways, which is beneficial for the used algorithm to place objects. There is also a high variety of different types of these nodes (for example river, mountain, forest, meadow), which can be found within a small area.

The duration of the field trial was set to two hours, whereby one hour was used for the evaluation methods and the other for playing the game. The starting time for the field trial execution was scheduled for 6 p.m. and lasted until 8 p.m. None of the participants were under time pressure.
The attendees were also given an optional goal to reach. They were told to reach level five with their character until the end of the game play session. After the session all of them were invited to free drinks in a nearby locality, where the questionnaires were filled out.

5.3.1 Data Collection in Field Trial

Although *Geo Heroes* is a client-side application which only uses *OpenStreetMap* to determine the location of game content, an additional server infrastructure was needed to log data to a server to record the emotions centralized. Therefore, a server-side backend platform powered by *Gamesparks*[1] was used to reduce the high workload. *Gamesparks* provides out-of-the-box features for server-side applications especially for multiplayer games. In *Geo Heroes* only the logging functionality was used to send data from a participant's device to the server-side interface during the field experiment. *Gamesparks* provides a section for server-side code called *Cloud Code*. Here the triggered events can be handled and processed with *Javascript* code. The server-side application was used to persist the received client data in a common database. Since *Gamesparks* uses a *mongoDB*, the log-data was transformed in json format to save it in a *mongoDB* collection.

5.3.2 Execution

Every participant was informed during registration and reminded a day before about place and time when the field trial starts. Nine out of ten came to the arranged appointment. Before starting the game play session, they were welcomed and introduced in the field trial's organization and execution. Since *Geo Heroes* only supported the operating system *Android* at this time, the participants used their own *Android* devices or the ones provided by the supervisor. Afterwards, the attendees were told to download and install the APK-file for using *Geo Heroes* on their smartphones. The participants were given a short introduction in starting and using the game but received not too much information, since the learning factor was also part of the evaluation. The field trial was split up in the following parts:

Personal Data Questionnaire Before the game-play session started, participants were told to fill in a personal data questionnaire. The questions

[1] http://www.gamesparks.com/

Participant	P1	P2	P3	P4	P5	P6	P7	P8	P9
$Version_{osm}$			•	•	•		•		•
$Version_{rnd}$	•	•				•		•	

Table 5.2: In this table the participants' installed and played versions are shown. $Version_{osm}$ stands for the version with the OpenStreetMap algorithm, whereas $Version_{rnd}$ is the one with the random algorithm to spawn game content.

Participant	P1	P2	P3	P4	P5	P6	P7	P8	P9
Hours/week	0	11	1	10	0	0	2	0	0
$Gamer_{yes}$		•	•	•			•		
$Gamer_{no}$	•				•	•		•	•

Table 5.3: Participants were divided into gamers and non-gamers depending on the played hours per week.

included for example the installed version, age, sex and experience in gaming. With this information it was possible to classify attendees into two categories. The first category was a division between the version of the game. The versions were distinguished by the implementation of the content spawning algorithm. *Version OSM* used the originally implemented *OpenStreetMap* algorithm, where each object was generated and placed with a relation to the real world. For example, a crab combatant could be found in the near of a river or sea. In contrast to *Version OSM*, *Version RND* just used a random algorithm for spawning. This means that every object could be found in every place. Table 5.2 shows the distribution of participants between the played versions.

The second category was defined by the personality of the participants. Since the 9 attendees consisted of five non-gamers and four gamers, a differentiation between them could be interesting and was therefore applied (Table 5.3). Thought was given to both categories, which were applied at one session and therefore gamers and non-gamers used both versions.

The questionnaire was created in *Google Forms* and results were automatically transformed into a *Google Spreadsheet*. In the spreadsheet, calculations

such as average, standard deviation and *t-test* were performed.
The advantage of using applications from *Google* is that it takes care of new participants. Those can join a new (supervised or unsupervised) field trial and calculations are updated automatically with the new entries.

Game Play Session The game play session took about one hour. The time span was chosen due to time constraints of the participants. Also, the game was in an early state and for the first iteration in an user-centered approach it is acceptable for such a low-fidelity prototype (Bowser et al., 2013). The participants were told to play the game in their own individual ways. There was just a rough overall goal to reach level five with their character. Since experience is the base for reaching levels and is received for nearly all interactions, attendees had great freedom on how to play the game. They were also told to integrate the game into their daily routine as long as it did not take too long.
In the first half hour participants were told to make use of their smartphone earphones. Afterwards they put them away and played the second half without them.
Due to the fact that the field trial should also record emotions, participants were presented a dialog containing 24 emotions displayed as buttons in a grid view after each interaction. By clicking on one of these "emotion-buttons" the game transfers log-data to the *Gamesparks* application. The data object, send in a *json* format, consisted of the following information:

- playerID
- timestamp
- type of interacted object
- name of interacted object
- longitude/latitude coordinates
- OpenStreetMap tag
- result of interacted object
- player emotion
- player level
- player life in percentage

The data object was then saved in a *mongoDB* collection as a document entry.

Each document therefore represented an emotion entry of a participant. An example document is shown in Listing 5.1.

Listing 5.1: Example of a logged data object

```
{
 "playerID": "P7",
 "currentDateTime": "June 15, 2016 4:25:14 PM UTC",
 "objectType": "ENEMY",
 "objectName": "Panther",
 "objectLon": "15.43888",
 "objectLat": "47.07441",
 "objectTag": "forest",
 "playerResult": "BATTLE_WIN",
 "playerEmotion": "Happy",
 "playerLevel": 2,
 "playerLife": 100
}
```

After the game play session, all participants met at a locality close to the field trial's starting point to receive a fare for their effort. They were told to quit *Geo Heroes* and to open the website where all remaining questionnaires were listed.

Game Experience Questionnaire & Pervasive Gameflow Questionnaire The GEQ was filled in immediately after game play. First, the *Core Module* was done, since it asks questions about feelings during game play. Second, participants filled in the *Post-Game Module*, where they were confronted with questions about their current feelings. Finally, the *Pervasive Game Flow Questionnaire* was completed by attendees.
Participants were accompanied by a supervisor while filling in the questionnaires. All questionnaires were created by *Google Forms* such as the *Personal Data Questionnaire* and results were transformed into *Google Spreadsheets* for further statistical analysis.

Open Discussion After each participant had filled in all questionnaires, a semi structured discussion was led. There were three general questions

prepared by the supervisor to create a flow of speech between all the persons present:

1. Have there been any problems while playing *Geo Hereoes*?
2. What did you not like about *Geo Heroes*?
3. What did you like most about *Geo Heroes*?

Everyone was free to speak and necessary information was filtered and recorded by the supervisor. Since everyone's native language was German, the discussion was also conducted in this language to eliminate language barriers.
When everything had been said, the discussion was brought to an end by the supervisor and the final lottery started. The lottery consisted of three 15 Euro *Amazon* vouchers and each of the participants had equal chances to win one of them. Afterwards the supervisor thanked everyone for participating and with these words the field trial ended.

6 Results

In this chapter, the results from the empirical study described in section 5.1 are presented. Since participants filled in a personal data questionnaire to find relations and categorize them into groups, this outcome is depicted first. Afterwards, the results of the *Game Experience Questionnaire (GEQ)* and *Pervasive Game Flow Questionnaire (PGFQ)* is depicted. Following this, the reported logged data of the player's emotion during the field experiment is visually presented in form of an emotion wheel based on Baillie et al. (2011). Finally, the major points of the open discussion are translated from German to English and transcribed for the reader. Since more than one evaluation method was conducted, the results are split up in their corresponding sections.

6.1 Personal Data Questionnaire

Before the game-play session of the field experiment started, the participants were told to fill in a *Personal Data Questionnaire (PDQ)*. The PDQ was primarily used for further statistical analysis. Besides asking about general information of the participants, the questions also include questions related to gaming, such as experience, favorite genres, platforms, modding and hours played per week. A full list of questions and answers is shown in Table 6.1. In question three, *What is your experience in gaming in general?*, it was asked about the overall experience by providing a *Likert* scale from zero, indicating that the participant never plays games, to 4, which represents a passionate gamer. In question four, *What are your favorite gaming genres?*, the participants reported their favorites in a multiple-choice selection. The genres, which can be viewed in the table, were described in more detail in

the original questionnaire, for example *Role-playing (Action, MMO, Roguelikes, Tactical, Sandbox, Fantasy)*. This is omitted in Table 6.1 due to space constraints and to retain a better overview in the table. This additional information was added to the choices for participants to better distinguish between the genres, as it is often not clear what the difference between two genres is. A possibly missing gaming genre in the list was covered by adding the option *Others*, which no one selected and it is therefore not presented in the provided table. The same *Other* choice is omitted in the presented table for question five, *On which platform are you playing games?*, since nobody has chosen this option.

6.2 Data Analysis

The *GEQ* and *PGFQ* were conducted with a categorization of the results from the *Personal Data Questionnaire (PDQ)*. In the PDQ, participants were asked to fill in the hours they play per week. When participants answered with zero hours, they were categorized as non-gamers (G_{no}). Participants who play at least one hour per week were classified as gamers (G_{yes}).
A further classification was introduced due to the installed and played game version of *Geo Heroes* during the field trial. One group was chosen to play the version with the *OpenStreetMap* spawning algorithm (V_{osm}). Here each game object has a relation to the environment of the physical world. For example, a *Pine Tree* will only spawn on a place where real trees in the world occur, such as in a forest or avenue. The other group of participants were given a version with a random spawning algorithm (V_{rnd}). This version spawns game content randomly to some place on the map without any relation to *OpenStreetMap* or other cartographic reference. The number of spawned objects was the same in both versions.
Each questionnaire consists of multiple components, and each component consists of multiple questions. In the following sections the results of GEQ and PGFQ are presented which includes the average for all participants (ϕ), for game versions and for gamers and non-gamers. Further statistical analysis is required to recognize a difference between the two categorizations. Therefore, standard deviation and an *independent two-sample t-test* analysis was performed to find statistical significance in the version ($t - test_V$) and

Nr	Questions	P1	P2	P3	P4	P5	P6	P7	P8	P9
1	Age	26	26	26	32	27	36	34	29	30
2	Sex	F	M	M	M	M	M	M	M	F
3.	What is your experience in gaming in general (0-4)?	0	2	4	2	1	1	3	0	0
4.	Favorite Gaming Genre									
	- Action		x		x					
	- Action-Adventure			x	x					
	- Adventure			x	x					
	- Role-playing				x			x		
	- Simulation		x					x	x	
	- Strategy				x	x		x	x	
	- Sport		x			x	x			
5.	On which platform are you playing games?									
	- PC		x	x	x	x	x	x	x	
	- Console		x	x	x	x				
	- Mobile			x			x			
6.	How many hours per week do you play on average?	0	11	1	10	0	0	2	0	0
7.	Are you interested in contributing in games (user-generated content, modding,...)	?	no	yes	yes	no	?	no	no	no
8.	What's your current status?	E	S	E	SE	S	E	U	S	E

Table 6.1: The table describes answers from the *Personal Data Questionnaire* for each participant. Question three was rated with a Likert scale between zero and four, where four stands for a very high experience in gaming. The question marks in question seven indicate that participants do not know whether they are interested in contributing to games or not. Question eight represents the current employment status: E = Employee, SE = Self Employed, S = Student, U = Unemployed

gamer category ($t - test_G$). If the result of the t-test between the two options of a category undercuts the threshold of 0.1, a tendency to significant difference is recognized. In the tables the cells of such values are colored *yellow*. If the value is lower than 0.05, significant difference is achieved and colored *green* in the tables representing the results. All questions were rated with a *Likert* scale between zero and four. The numbers represent the 'agree-factor' of a participant to the related question. Zero means that the participant did *not agree at all* and four means an *extreme* agreement to the provided question.

T-test Calculation in Components
An *independent two-sample t-test* calculation was carried out for each question and component. The participants were divided into an experimental (exp) and a control group (con). The compilation of groups was defined by version and game experience classification. Since existing questionnaires and defined statements were used which described the components, the 'per component' *t-test* was carried out with the group's average values of each question in the component:

$$T_{test} = (\{Avg_{1_{exp}}, Avg_{2_{exp}}, ...\}, \{Avg_{1_{con}}, Avg_{2_{con}}, ...\})$$

$Avg_{1_{exp}}$ is the average of all answers in the first question in the component from the experimental group.

6.3 Game Experience Questionnaire

The results of the *Game Experience Questionnaire* are presented in this section. Since this questionnaire relies on two modules, the *Core Module* and a *Post-Game Module*, results are divided into paragraphs.

Experience Questionnaire - Core Module In the *Core Module* of the *Game Experience Questionnaire* seven components are analyzed in the course of this empirical evaluation. The results of the components are shown in Table 6.2.

Component	ϕ	V_{osm}	V_{rnd}	G_{yes}	G_{no}	$t-test_G$	$t-test_V$
Competence	1.64	2.00	1.20	2.05	1.32	0.03	0.04
Sensory and Imaginative Immersion	1.67	1.80	1.50	1.50	1.80	0.38	0.37
Flow	2.98	3.00	2.95	2.85	3.08	0.30	0.8
Tension / Annoyance	1.19	1.00	1.42	1.83	0.67	0.03	0.23
Challenge	1.67	1.56	1.80	1.85	1.52	0.40	0.46
Negative affect	1.50	1.50	1.50	1.56	1.45	0.84	1.00
Positive affect	2.36	2.44	2.25	2.20	2.48	0.26	0.45

Table 6.2: Results for each component in the Core module of the *Game Experience Questionnaire* are listed in this table. Green highlighted cells indicate a significant difference in the t-test results.

A visual representation of these values is shown in Figure 6.1. The t-test values of the *Competence* component are 0.04 in case of the played version and 0.03 for the gamer category, which implies that there is a significant difference between the elements in these categories. Figure 6.2 shows the variation in the categories. The *OpenStreetMap* version has a value of 2.00, whereas the *Random* version has just 1.20. Also, gamers rated this component higher (2.05) than non-gamers (1.32). The *Tension / Annoyance* component shows also a statistical significance in the *t-test* (0.03) of the gamer category (Figure 6.3). The gamers' value in this component is much higher (1.83) than the non-gamers' value (0.67).
Only the components *Competence* and *Tension/Annoyance* have shown a significant difference with the t-test method. Other components are therefore described with the mean value of all participants. *Sensory and Imaginative Immersion* has a mean of 1.67. *Flow* is measured with a value of 2.98 on average. The average value of *Challenge* is 1.67. *Positive affect* is measured with 2.36 and the *Negative affect* has a value of 1.50. In Table 6.4 each single question is listed with the appropriate results. Every question relates to just one component. For the *GEQ Core Module* the mapping is listed in Table 6.3. Since each component consists of multiple questions a 'per-question' t-test analysis was also conducted. The majority of questions does not have significant differences between the results in the categories. However, some

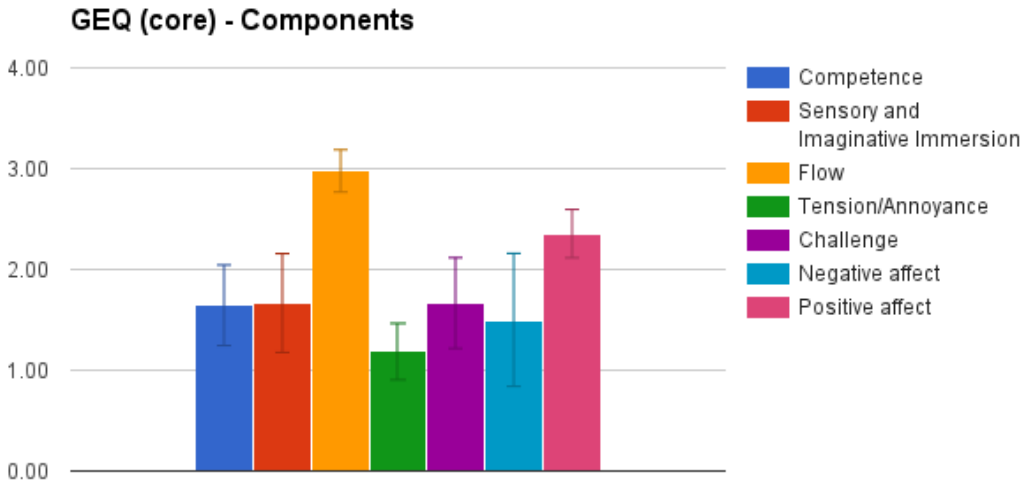


Figure 6.1: The results of components of the Core-module in the Game Experience Questionnaire presented in a bar chart.

Component	Question IDs
Competence	2,10,15,17,21
Sensory and Imaginative Immersion	3,12,18,19,27,30
Flow	5,13,25,28,31
Tension/Annoyance	22,24,29
Challenge	11,23,26,32,33
Negative affect	7,8,9,16
Positive affect	1,4,6,14,20

Table 6.3: The table maps the given questions represented by their IDs to the related components of the *Game Experience Questionnaire - Core Module*

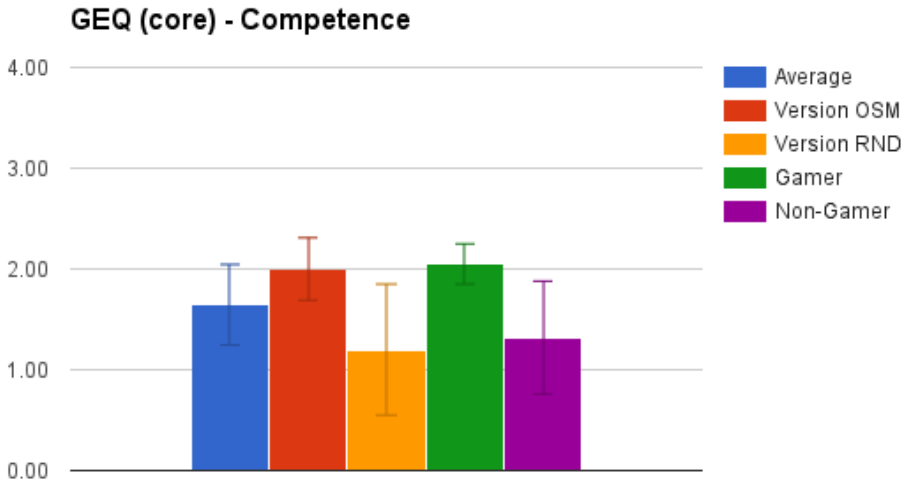


Figure 6.2: The component *Competence* of the Game Experience Questionnaire - Core module is shown in this figure. Each bar represents a group of participants. Played version and gamer category showed both a significant difference.

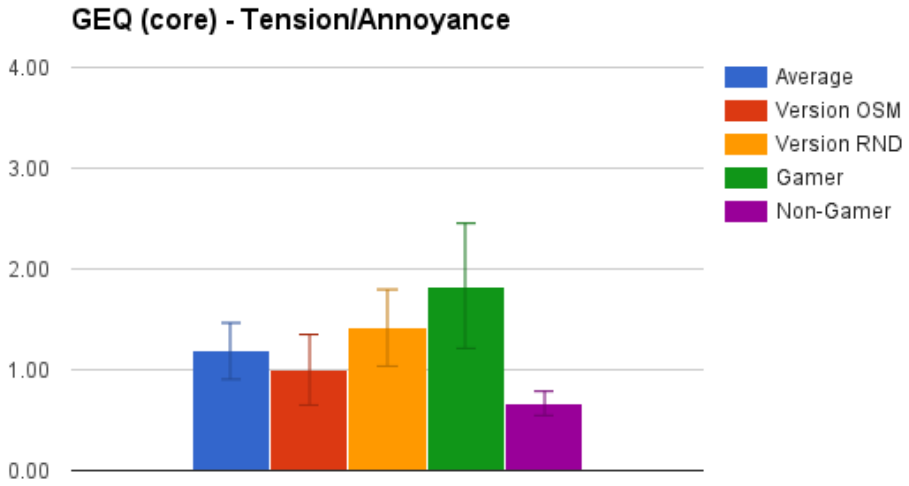


Figure 6.3: The component *Tension/Annoyance* of the Game Experience Questionnaire - Core module is shown in this figure. Each bar represents a group of participants. Gamers were much more annoyed than non-gamers.

of them vary significantly. First, the results for the played versions are reported. Question 6 - *"I felt happy"* had a value of 2.6 in the *OpenStreetMap* version, whereas in the *Random* version the value was just 1.5. Question 11 - *"I thought the game was hard"* was evaluated with 0.8 in the *OpenStreetMap* version, but 1.75 in the *Random* one. Furthermore, question 21 - *"I was fast at reaching the game's target"* was easier in the *OpenStreetMap* version (2.0) than in the *Random* version (0.5) and therefore showed a tendency to significance. In the gamer and non-gamer category significant differences were also found. In question 14 - *"I felt good"*, non-gamers reported on average a much higher value (2.6) than gamers (1.5) and showed a tendency to significance. In question 29 - *"I felt frustrated"*, significance was measured. Here, gamers had a high value of 2.5, whereas non-gamers just had a value of 0.6.

Nr.	Question	ϕ	V_{osm}	V_{rnd}	G_{yes}	G_{no}	$t-test_G$	$t-test_V$
1	I felt content	2.44	2.40	2.50	2.50	2.40	0.85	0.85
2	I felt skillful	1.22	1.8	0.5	1.75	0.8	0.26	0.11
3	I was interested in the game's story	1.11	1.20	1	1	1.20	0.80	0.80
4	I thought the game was fun	2.67	2.60	2.75	2.50	2.80	0.41	0.68
5	I was fully occupied with the game	3.11	3.20	3.00	3.00	3.20	0.73	0.73
6	I felt happy	2.11	2.60	1.50	2.25	2.00	0.66	0.02
7	Playing Geo Heroes put me in a bad mood	0.89	0.60	1.25	1.00	0.80	0.83	0.48
8	I thought about other things	1.89	2.40	1.25	1.50	2.20	0.48	0.23
9	I found the game tiresome	2.22	2.00	2.50	2.75	1.80	0.21	0.53
10	I felt competent	1.56	1.60	1.50	2.00	1.20	0.37	0.91
11	I thought the game was hard	1.22	0.80	1.75	1.00	1.40	0.51	0.09
12	The game was aesthetically pleasing	1.89	1.6	2.25	1.50	2.20	0.29	0.33
13	I forgot everything around me	3.00	3.20	2.75	2.75	3.20	0.38	0.38
14	I felt good	2.11	2.00	2.25	1.50	2.60	0.07	0.72
15	I was good at playing the game	2.00	2.20	1.75	2.25	1.80	0.68	0.68
16	I felt bored	1.00	1.00	1.00	1.00	1.00	1.00	1.00
17	I felt successful	2.11	2.40	1.75	2.25	2.00	0.66	0.24
18	I felt imaginative	1.00	1.20	0.75	1.50	0.60	0.20	0.54
19	I felt that I could explore things	1.89	2.20	1.50	1.50	2.20	0.36	0.36
20	I enjoyed the game	2.44	2.60	2.25	2.25	2.60	0.51	0.51

21	I was fast at reaching the game's targets	1.33	2.00	0.50	2.00	0.80	0.15	0.06
22	I felt annoyed	1.22	0.80	1.75	1.75	0.80	0.34	0.34
23	I felt pressured	1.33	1.20	1.50	2.00	0.80	0.22	0.77
24	I felt irritable	0.89	0.80	1.00	1.25	0.60	0.48	0.83
25	I lost track of time	3.22	3.00	3.50	3.50	3.50	0.48	0.48
26	I felt challenged	2.33	2.40	2.25	2.25	2.40	0.87	0.87
27	I found the game impressive	2.22	2.60	1.75	2.00	2.40	0.51	0.14
28	I was deeply concentrated on the game	2.67	2.80	2.50	2.25	3.00	0.35	0.72
29	I felt frustrated	1.44	1.40	1.50	2.50	0.60	0.03	0.92
30	The game felt like a rich experience	1.89	2.00	1.75	1.50	2.20	0.20	0.66
31	I lost connection with the outside world	2.89	2.80	3.00	2.75	3.00	0.57	0.65
32	I felt time pressure	1.89	2.00	1.75	2.50	1.40	0.29	0.82
33	I had to put a lot of effort into the game	1.56	1.40	1.75	1.50	1.60	0.89	0.64

Table 6.4: Results for each question in the Core module of the *Game Experience Questionnaire* are listed in this table. Green highlighted cells indicate a significant difference in the t-test results. Yellow highlighted cells indicate a tendency to significant difference in the t-test results.

Component	ϕ	V_{osm}	V_{rnd}	G_{yes}	G_{no}	$t-test_G$	$t-test_V$
Positive Experience	1.31	1.63	0.92	1.46	1.20	0.46	0.09
Negative Experience	0.81	0.67	1.00	1.00	0.67	0.39	0.41
Tiredness	1.72	1.40	2.13	2.00	1.50	0.20	0.09
Returning to Reality	0.89	1.00	0.75	0.83	0.93	0.89	0.75

Table 6.5: Results for each component in the Post-game module of the *Game Experience Questionnaire* are listed in this table. Yellow highlighted cells indicate a tendency to significant difference in the t-test results.

Game Experience Questionnaire - Post-Game Module The *Post-Game Module* consists of four components, which are statistically analyzed for this field experiment. A graphical overview of all four elements is shown in Figure 6.4. For a detailed analysis of the measured components Table 6.5 presents the values including t-test to illustrate significance in the gamer and version categories. The average value of the *Positive Experience* component receives a value of just 1.31, and the *Negative Experience* is measured with a value of 0.81. However, the *Positive Experience* shows a tendency of significance with a value of 0.09 in relation to the version. The *OpenStreetMap* version has a value of 1.63, whereas the *Random* version has just 0.92 (Figure 6.5).
The value in the *Tiredness* component is 1.72 on average. The t-test for versions shows a tendency of significance with a value of 0.09. The value in the *Random* version is much higher (2.13), than the one in the *OpenStreetMap* version (1.40). A visual representation of the *Tiredness* component is shown in Figure 6.6. The *Returning to Reality* component was measured with a value of 0.89 on average, but does not show a significant difference in the categories. The affiliation of question IDs to the components is shown in Table 6.6.

As already described in the *Core Module*, a 'per question' analysis for the *Post-Game Module* was conducted including a t-test for gamers and versions. A full list of the 'per question' analysis is shown in Table 6.7. In question five, *"It felt like a victory"*, the different versions show a tendency to sig-

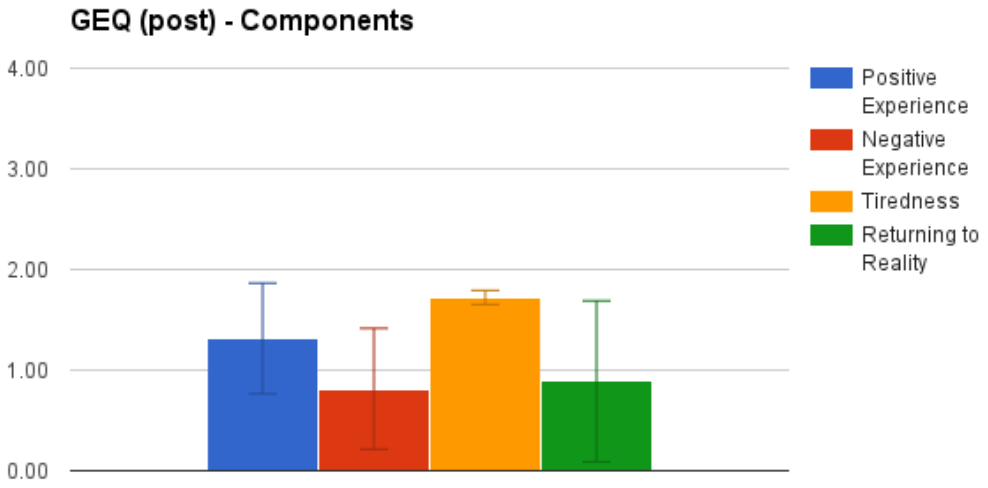


Figure 6.4: The results of components of the Post-game module in the Game Experience Questionnaire presented in a bar chart.

Component	Question IDs
Positive Experience	1,5,7,8,12,16
Negative Experience	2,4,6,11,14,15
Tiredness	10,13
Returning to Reality	3,9,17

Table 6.6: The table maps the given questions represented by their IDs to the related components of the *Game Experience Questionnaire - Post-Game Module*

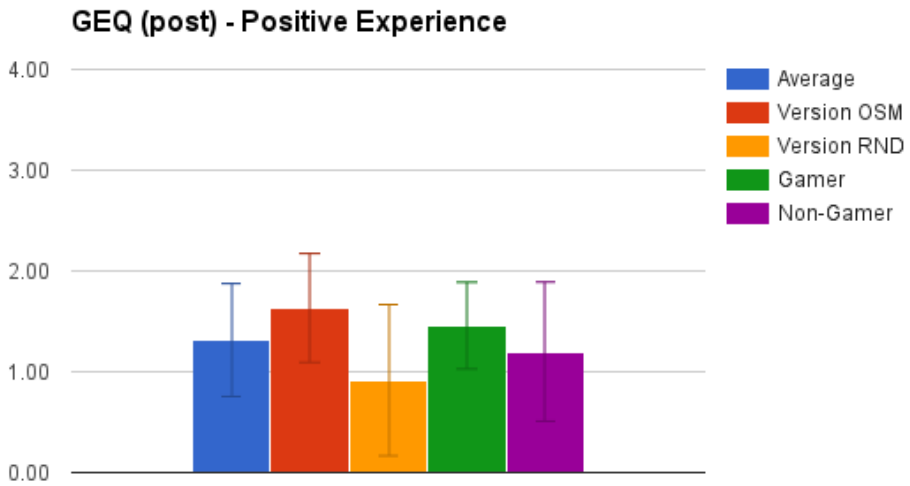


Figure 6.5: The component *Positive Experience* of the Game Experience Questionnaire - Post-game module is shown in this figure. Each bar represents a group of participants. Statistical significance is shown between the *OSM* and *Random* version.

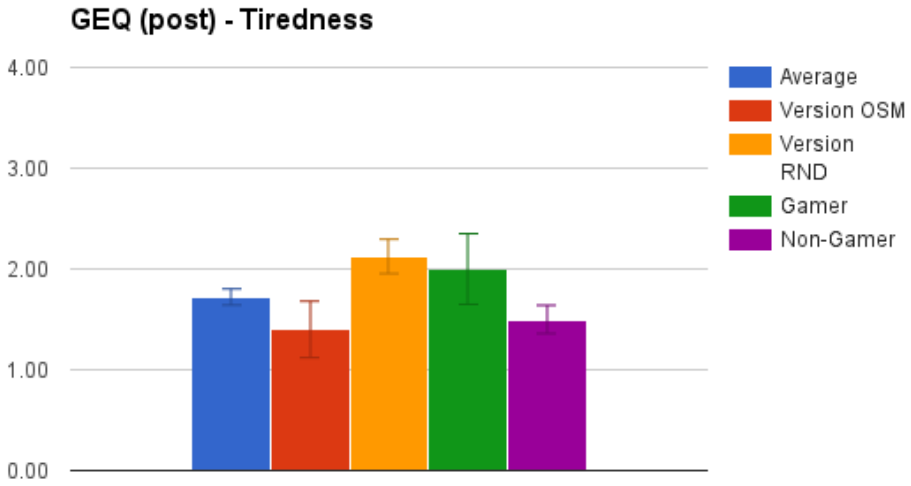


Figure 6.6: The component *Tiredness* of the Game Experience Questionnaire - Post-game module is shown in this figure. Each bar represents a group of participants. Statistical significance is shown between the *OSM* and *Random* version.

nificance with a value of 0.09. The *OSM* version received a value of 2.00, whereas the *Random* version scored 1.5. Question twelve, *"I felt powerful"*, was rated with only 1.22 on average for all participants. Here the t-test for game versions shows a significant difference with a value of 0.02. The *OSM* version received a value of 2.00, whereas the *Random* version only got 0.25. Question 14, *"I felt regret"*, also shows a tendency to significance with a value of 0.09 in the played version category. Even if both versions scored very low, the random version has a value of 0.5, whereas the OSM version scored 0.00 in this case. In question 16, *"I felt proud"*, the version and the gamer category shows significance. The t-test for the gamer category was measured with 0.01, and for the version category the t-test reveals a value of 0.07 which indicates at least a tendency to significance. It should be noted, that the average value for this question was really low. The average of all participants received just 0.33. The *OSM* version had 0.60 and the *Random* version was 0.00. The difference for gamers was higher, but also here, gamers only gained 0.75 and non-gamers 0.00. In the seventeenth question, *"I had the feeling that I had returned from a journey"* participants reported on average just 1.78. But the measurement for the game versions showed a significant difference of 0.05. The *OSM* version received a total score of 2.40. In contrast, the *Random* version just gained 1.00.

Nr.	Question	ϕ	V_{osm}	V_{rnd}	G_{yes}	G_{no}	$t-test_G$	$t-test_V$
1	I felt revived	1.11	1.6	0.5	1.25	1	0.75	0.12
2	I felt bad	0.67	0.40	1.00	0.75	0.60	0.84	0.41
3	I found it hard to get back to reality	0.22	0.20	0.25	0.25	0.20	0.88	0.88
4	I felt guilty	0.56	0.60	0.50	1.00	0.20	0.10	0.85
5	It felt like a victory	1.78	2.00	1.50	1.75	1.80	0.88	0.09
6	I found the game a waste of time	1.11	0.60	1.75	1.00	1.20	0.83	0.19
7	I felt energized	1.67	1.60	1.75	1.50	1.80	0.68	0.84
8	I felt satisfied	1.78	2.00	1.50	2.00	1.60	0.41	0.29
9	I felt disoriented	0.67	0.40	1.00	0.50	0.80	0.64	0.33
10	I felt exhausted	1.78	1.6	2.00	2.25	1.40	0.40	0.70
11	I felt that I could have done more useful things	1.89	1.80	2.00	2.00	1.80	0.87	0.87
12	I felt powerful	1.22	2.00	0.25	1.50	1.00	0.57	0.02
13	I felt weary	1.67	1.20	2.25	1.75	1.60	0.84	0.12
14	I felt regret	0.22	0.00	0.50	0.25	0.20	0.88	0.09
15	I felt ashamed	0.44	0.60	0.25	1.00	0.00	0.15	0.64
16	I felt proud	0.33	0.60	0.00	0.75	0.00	0.01	0.07
17	I had the feeling that I had returned from a journey	1.78	2.40	1.00	1.75	1.80	0.95	0.046

Table 6.7: Results for each component in the Post-game module of the *Game Experience Questionnaire* are listed in this table. Green highlighted cells indicate a significant difference in the t-test results. Yellow highlighted cells indicate a tendency to significant difference in the t-test results.

6.4 Pervasive Game Flow Questionnaire

After the *GEQ* (core and post-game module), participants filled in the *PGFQ*. It should therefore be considered, that there was a time delay of about ten to fifteen minutes between game session and the questions of the PGFQ. The PGFQ consists of 34 questions which are assigned to seven individual components. Each question is a *Likert* scale question with the range of zero, indicating disagreement, to four, for total agreement. A t-test was applied to all components and to all questions. First of all, the results of the components are described.
From seven components only *Player Skills* and *Concentration* show a significant difference or tendency in the evaluation. For *Player Skills*, the t-test for versions and for gamers revealed statistical significance or a tendency to it. The t-test for the gamer category in this component has a value of 0.06 and the t-test for version scored 0.04. The average of the *OSM* version is just 1.94, whereas the *Random* version received 2.50. The non-gamers' average is 2.49 and the value for participants with gaming experience is only 1.82 (Figure 6.7). The t-test for the *Concentration* component scored a value for gamers of 0.095. In this case, gamers scored an average of 1.63, whereas non-gamers rated the questions in this component with an average value of 2.33 (Figure 6.8). However, other components do not show any indication of a statistical significance and are therefore described with their average of all participants. *Challenge* was rated with 1.89 and *Control* with 2.10. *Clear Goals* only received 1.39 in the evaluation, and *Feedback* got 2.37. The last component was *Immersion* and scored a value of 1.96. An overview of the averages of all components including standard deviation is shown in Figure 6.9. The higher the value of a component, the better the game performs in the evaluation. Therefore, an average was built upon all components. The value of the combined score received 1.99 points. An overview of results of each component is shown in Table 6.8. The questions in the PGFQ are described in more detail than the questions in the GEQ. A per-question analysis can therefore be important for further research on how to improve the game. Hence, a t-test is applied for each question separately. However, all but three of the questions do not reveal a significance in the categories. The question *"I was just confronted with tasks that felt important"* from the *Concentration* component was rated by participants with an average of 1.78. The

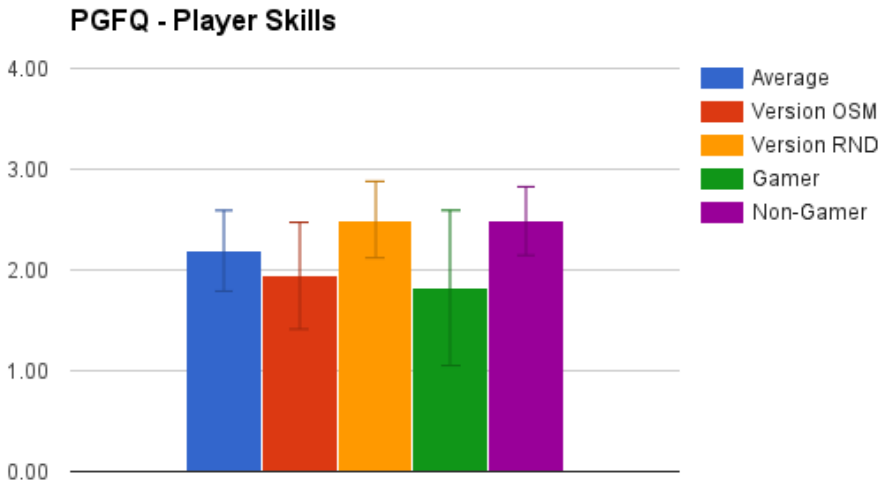


Figure 6.7: The results of the component *Player Skills* of the Pervasive Game Flow Questionnaire presented in a bar chart.

Component	ϕ	V_{osm}	V_{rnd}	G_{yes}	G_{no}	$t-test_G$	$t-test_V$
Concentration	2.02	1.97	2.08	1.63	2.33	0.095	0.79
Challenge	1.89	1.85	1.94	1.69	2.05	0.54	0.88
Player Skills	2.19	1.94	2.50	1.82	2.49	0.06	0.04
Control	2.10	2.11	2.07	1.89	2.26	0.26	0.87
Clear Goals	1.39	1.40	1.38	1.25	1.50	0.45	0.97
Feedback	2.37	2.53	2.15	2.33	2.40	0.88	0.48
Immersion	1.96	2.12	1.73	1.90	2.02	0.72	0.37

Table 6.8: Results for each component of the *Pervasive Game Flow Questionnaire* are listed in this table. Green highlighted cells indicate a significant difference in the t-test results. Yellow highlighted cells indicate a tendency to significant difference in the t-test results.

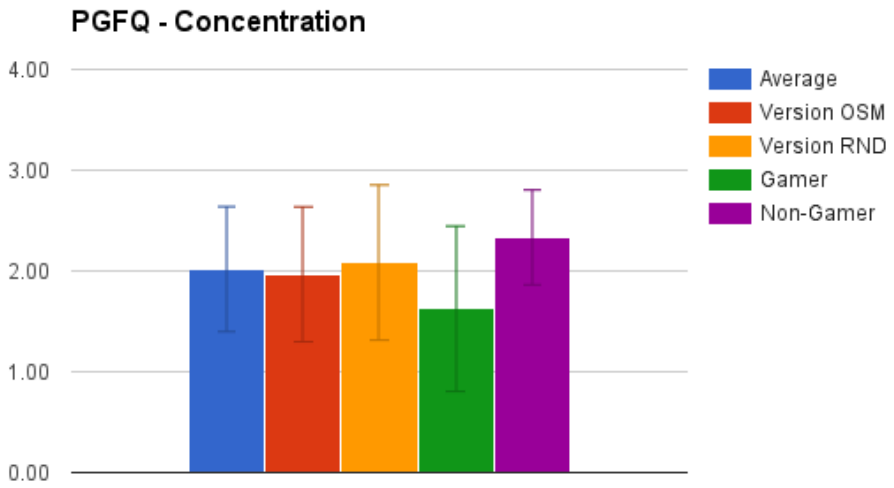


Figure 6.8: The results of the component *Concentration* of the Pervasive Game Flow Questionnaire presented in a bar chart.

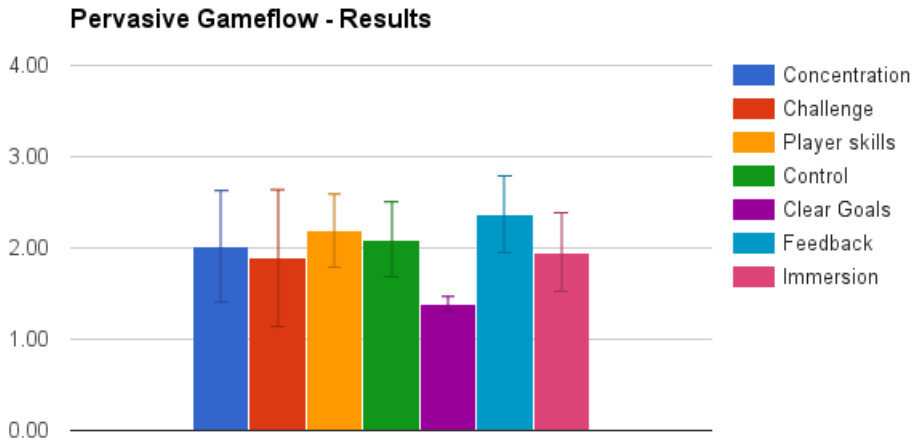


Figure 6.9: The results of all components of the Pervasive Game Flow Questionnaire presented in a bar chart.

OSM version only received a score of 1.20, whereas the *Random* version got a value of 2.50. The t-test of versions for this question scored 0.01. Gamers also rated this question really low with 1.25, in contrast to the non-gamers with 2.20. The t-test resulted in 0.05 for the gamer category. In the *Feedback* component, the question *"I received immediate feedback on my actions"* reveals also a significant difference in the version and the gamer group. The average got a value of 2.67. The *OSM* version was rated with 3.20 and the *Random* version with 2.00. T-test has shown a significant difference of 0.03 in the version category. The t-test for the gamer category resulted in 0.09. The average difference between these groups is small. Gamers rated this question with 2.75, whereas non-gamers rated it with 2.60. Statistical significance was also measured in the last question of the component *Immersion*. The t-test analysis of the question *"The game enabled me to shift focus between the virtual and physical parts of the game world without losing too much of the feeling of immersion"* revealed a significance with a value of 0.03 in the version category. The average score of all participants is 2.00. The *OSM* version was rated with 2.80, which is much higher than the Random version with 1.00.

An overview of results of each question is shown in Table 6.9. Since each question belongs to one component, Table 6.10 shows the mapping between components and questions of the PGFQ.

Nr.	Question	ϕ	V_{osm}	V_{rnd}	G_{yes}	G_{no}	$t-test_G$	$t-test_V$
1	The game provided a lot of stimuli from different sources	1.89	2.00	1.75	1.50	2.20	0.81	0.77
2	The game provided stimuli that are worth paying attention to	2.11	2.00	2.25	1.75	2.40	0.61	0.77
3	The game quickly grabbed my attention and maintained my focus throughout the game	3.11	3.20	3.00	3.00	3.20	1.00	0.80
4	I was just confronted with tasks that felt important	1.78	1.20	2.50	1.25	2.20	0.046	0.01
5	The game had a high workload, while still being appropriate for my perceptual, cognitive, and memory limits.	2.00	1.80	2.25	1.75	2.20	0.47	0.48
6	The game supported me in the process of switching concentration between in-game tasks and surrounding factors of importance (e.g. traffic)	1.22	1.60	0.75	0.50	1.80	0.87	0.43
7	The challenges in the game matched my skill level	2.22	2.20	2.25	1.75	2.60	0.67	0.96
8	The game provided different levels of challenge	2.44	2.40	2.50	2.50	2.40	1.00	0.91
9	The game stimulated and supported me in my own creation of game scenarios and pacing	0.78	0.80	0.75	0.50	1.00	0.74	0.94

10	The game helped me in keeping a balance in the creation of paths and developments in the game world, but didn't put too much control or constraints on the pacing and challenges evolving.	2.11	2.00	2.25	2.00	2.20	0.74	0.72
11	I was able to start playing the game without further information (e.g. tutorials, online help,...)	2.78	2.60	3.00	3.25	2.40	0.85	0.70
12	Learning the game was fun	2.44	2.00	3.00	1.75	3.00	0.15	0.15
13	The game increased my skills while progressing through the game	2.11	2.00	2.25	1.25	2.80	0.31	0.77
14	The game provided appropriate rewards for my effort	2.56	2.60	2.50	2.50	2.60	1.00	0.80
15	The game interfaces and mechanics were easy to learn	1.78	1.60	2.00	1.50	2.00	0.72	0.70
16	The game interfaces and mechanics were easy to use	1.78	1.20	2.50	1.25	2.20	0.36	0.24
17	The game was very flexible and enabled my skills to be developed in a pace set by myself.	1.89	1.60	2.25	1.25	2.40	0.20	0.33
18	I felt a sense of control over my character, movements and interactions in the game world	1.89	2.00	1.75	1.00	2.60	0.45	0.79
19	I felt a sense of control over the game interface and input devices	1.67	2.00	1.25	1.75	1.60	0.61	0.35

20	I felt a sense of control over the game shell (starting, stopping, saving, etc.)	2.33	2.20	2.50	2.25	2.40	0.82	0.74
21	I was not able to make errors that are detrimental to the game. (e.g. the game supports in recovering from errors)	1.89	1.80	2.00	2.25	1.60	0.80	0.80
22	My actions had impact on the game world.	1.89	1.80	2.00	1.25	2.40	0.42	0.82
23	I was free to play the game the way I wanted.	2.89	3.20	2.50	3.00	2.80	0.53	0.29
24	The game enabled me to easily pick up game play in a constantly ongoing game and quickly get a picture of the current status in the game world.	2.11	1.80	2.50	1.75	2.40	0.22	0.20
25	Overriding (Primary) goals were clear and presented early.	1.33	1.00	1.75	1.00	1.60	0.27	0.22
26	The game supported me in forming and communicating my own intermediate goals.	1.44	1.80	1.00	1.50	1.40	0.54	0.27
27	I received feedback on progress toward my goals.	2.56	2.80	2.25	2.75	2.40	0.54	0.46
28	I received immediate feedback on my actions.	2.67	3.20	2.00	2.75	2.60	0.09	0.03
29	I always knew my status and score	1.89	1.60	2.20	1.50	2.20	0.46	0.44
30	I became less self-aware and less worried about everyday life or self	2.22	2.00	2.50	2.25	2.20	0.82	0.60

31	I experienced an altered sense of time	2.33	2.20	2.50	2.25	2.40	0.78	0.72
32	I felt emotionally involved in the game	1.22	1.40	1.00	1.25	1.20	0.72	0.51
33	The game supported a seamless transition between different everyday contexts, and did not imply or require actions that might result in a violation of social norms in all contexts	2.00	2.20	1.67	1.50	2.50	0.74	0.58
34	The game enabled me to shift focus between the virtual and physical parts of the game world without losing too much of the feeling of immersion.	2.00	2.80	1.00	2.25	1.80	0.29	0.03

Table 6.9: Results for each questions of the *Pervasive Game Flow Questionnaire* are listed in this table. Green highlighted cells indicate a significant difference in the t-test results. Yellow highlighted cells indicate a tendency to significant difference in the t-test results.

Component	Question IDs
Concentration	1,2,3,4,5,6
Challenge	7,8,9,10
Player Skills	11,12,13,14,15,16,17
Control	18,19,20,21,22,23,24
Clear Goals	25,26
Feedback	27,28,29
Immersion	30,31,32,33,34

Table 6.10: The table shows the correlation between questions represented by their IDs to the related components of the *Pervasive Game Flow Questionnaire*.

6.5 Emotion Wheel

During the game play session, participants were forced to report their emotions after an interaction in *Geo Heroes*. A dialog with 24 different emotions in the form of buttons popped up and participants had to choose the one that described their current state best. Afterwards, the dialog disappeared and users could proceed with the game. The visual representations of emotions are based on the *Emotion Wheel* by Baillie et al. (2011). The diameter of the circle behind the name of the emotion indicates how often this entry was picked by participants. A small circle without the name of an emotion indicates that no participant has selected this emotion during the field experiment. In total 157 emotions from all participants were recorded and transferred to the server. The *Emotion Wheel* for the whole field trial is depicted in Figure 6.10. The chosen emotions were transferred to the *Gamesparks* server and saved in the database. Additionally, the outcomes of interactions were also sent to the server. Outcomes differed depending on the type of interaction, for example a win or loss in a battle. The database entries were then analyzed in their outcomes and also in the gamer/non-gamer and game version category. An example of a full log-entry was already depicted in Table 5.2. Since each emotion belongs to one of the four quadrants, active-positive, active-negative, passive-positive and passive-negative, the number of emotions was summed up for each axis and expressed as the percentage of the samples (Table 6.11).

Category	Emotion Axis			
	negative	positive	active	passive
OVERALL	22%	78%	32%	68%
ITEM_SUCCESS	6%	94%	43%	57%
BATTLE	24%	76%	35%	65%
SHOP	26%	74%	15%	85%
ITEM_MAX	8%	92%	58%	42%
ITEM_MED	0%	100%	31%	69%
ITEM_MIN	10%	90%	40%	60%
ITEM_NONE	60%	40%	40%	60%
BATTLE_WIN	11%	89%	36%	64%
BATTLE_LOSE	86%	14%	29%	71%
BATTLE_WIN_SPECIAL	6%	94%	43%	57%
SHOP_BOUGHT	8%	92%	8%	92%
SHOP_NOT_BOUGHT	35%	65%	19%	81%
VERSION_OSM	19%	81%	35%	65%
VERSION_RANDOM	24%	76%	28%	72%
GAMER	28%	72%	33%	67%
NON_GAMER	17%	83%	31%	69%

Table 6.11: The table shows the results of categories and different outcomes of interactions which were tracked and measured for the *Emotion Wheel*

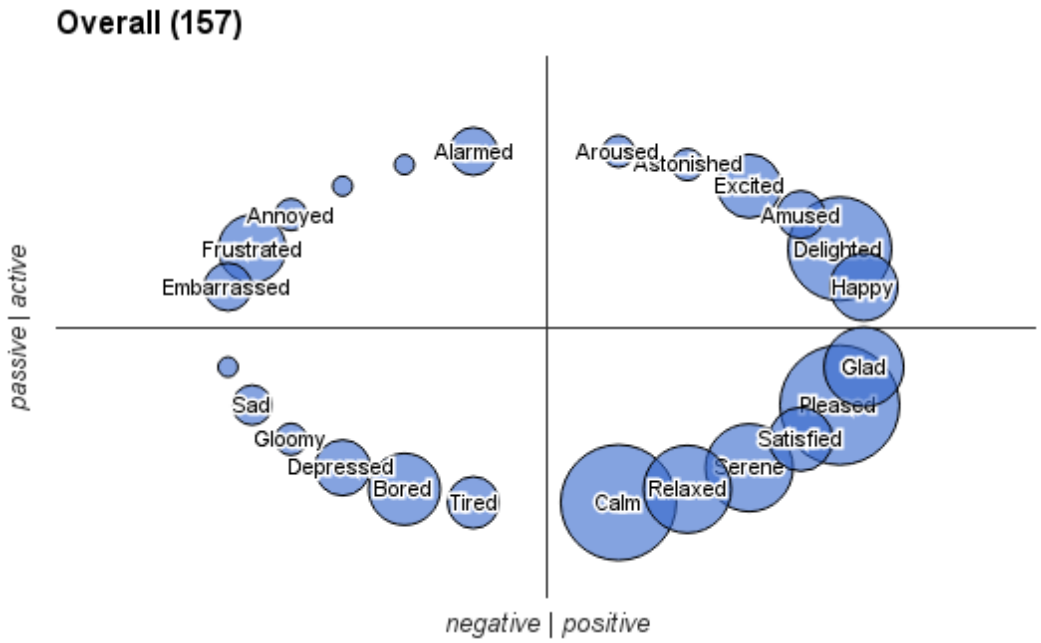


Figure 6.10: An overview of the results for the tracked emotions. 157 emotions were tracked. The larger the diameter of a circle behind an emotion, the more often participants chose that particular emotion. A small circle without a label indicates that no participant picked this emotion.

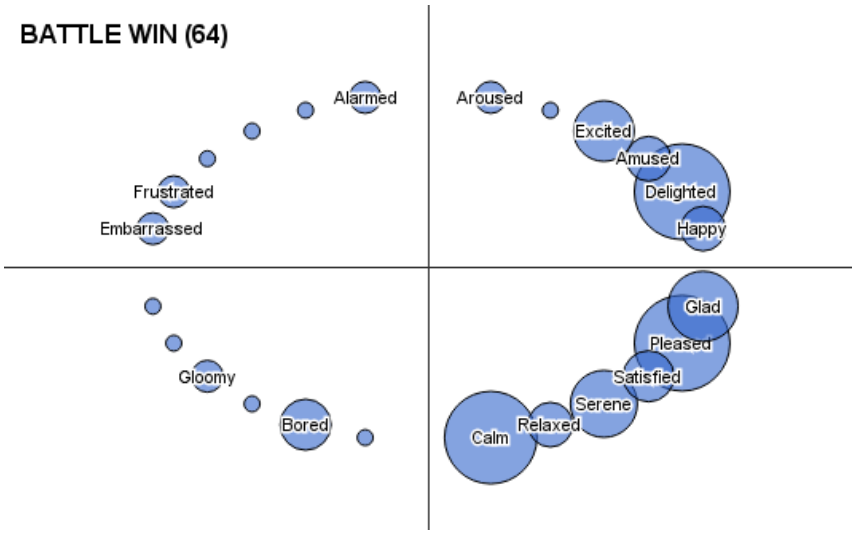


Figure 6.11: The logged emotion when players have won a battle. As the figure shows, active as well as passive positive emotions are picked most of the time.

Battle

The outcome of a battle was BATTLE_WIN when a player had won a battle, BATTLE_LOSE, when he or she had lost a battle, or BATTLE_WIN_SPECIAL, when a player had won the battle and received a random item as an additional reward. The highest difference is shown between the positive and negative quadrants. When participants won a battle, 89% of them had positive emotions (Figure 6.11), in contrast to a loss where only 14% were positive (Figure 6.12). A rise of positive emotions to 94% was recognized with the outcome BATTLE_WIN_SPECIAL.

Resources

Collecting a resource also had different outcomes. Depending on the probability of the amount of pieces a resource deposit contained, the outcome varied. NOTHING represented a collection of zero pieces of this resource. MINIMUM was used when getting exactly one piece. MEDIUM was chosen when the resource deposit provided two pieces and MAXIMUM when three pieces of a resource were collected. For further analysis, the sum of a successful collection of resources, which is the sum of logged emotions

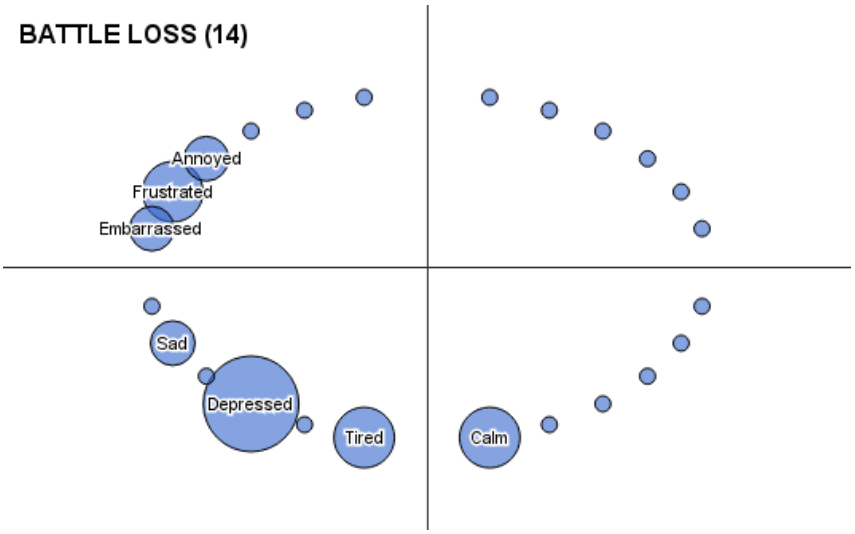


Figure 6.12: The logged emotion when players lost a battle. It is obvious, that players have negative emotions when not succeeding.

with outcomes MINIMUM, MEDIUM and MAXIMUM were combined and confronted (Figure 6.13) with the not successful collections represented by the outcome NOTHING (Figure 6.14). In case of a successful collection positive emotions were recorded with 94%. Even when participants were non-successful in collecting the resource, they were still passive-positively evolved with 40%. It should be mentioned that due to the probability of collecting resources, the number of samples for not successful collections was just five, whereas 35 samples were logged with successful collections.

Shop

Shop interactions were also divided into two outcome states. BOUGHT was used when a player had bought or sold at least one item, whereas NOT_BOUGHT was chosen when the shop was visited but nothing was consumed. In total there were 39 interactions with shops, of which 13 reported a buy or sell process by participants (Figure 6.15). However, when participants did not buy anything (Figure 6.16), the *Frustration* emotion was reported four times. The negative experience when nothing was bought was much higher with 35%, in contrast to a participant who made a transaction

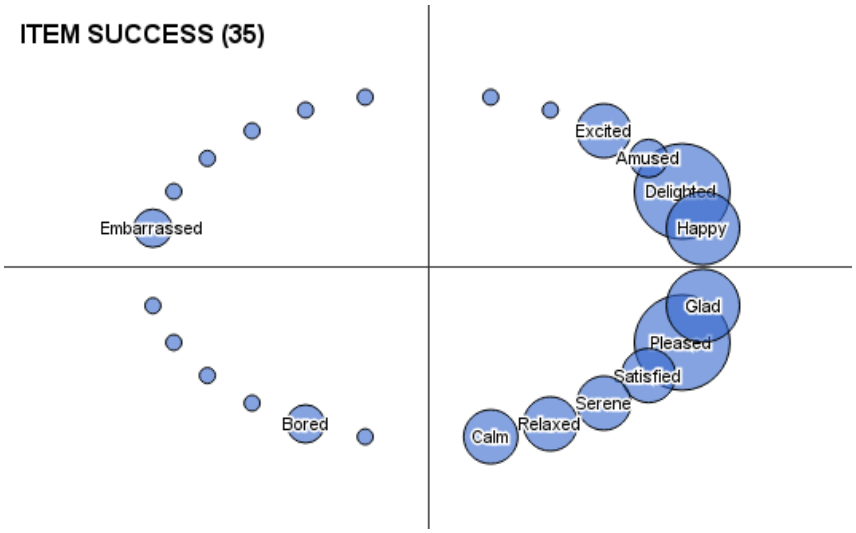


Figure 6.13: If an item was successfully collected, the emotions were almost always positive

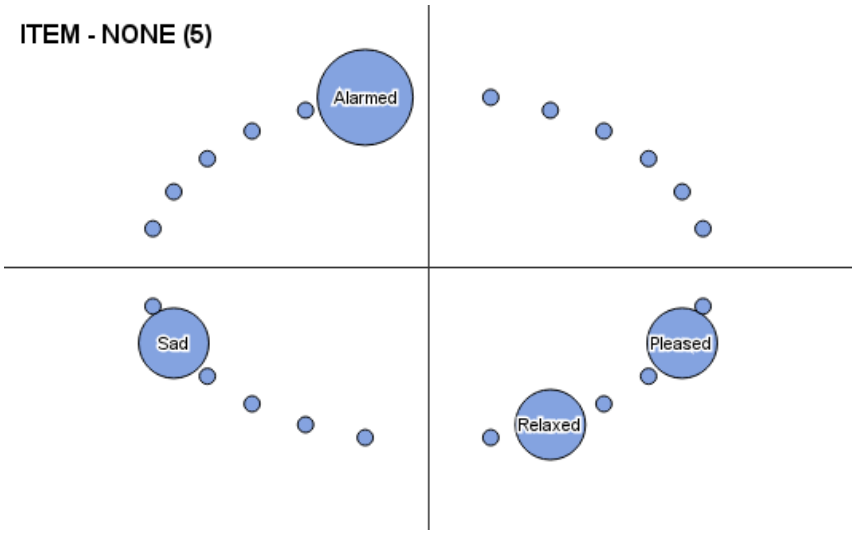


Figure 6.14: In case an item was not successfully collected, players had different emotional feelings about the current situation. It should be noted, that this interaction outcome has just occured five times.

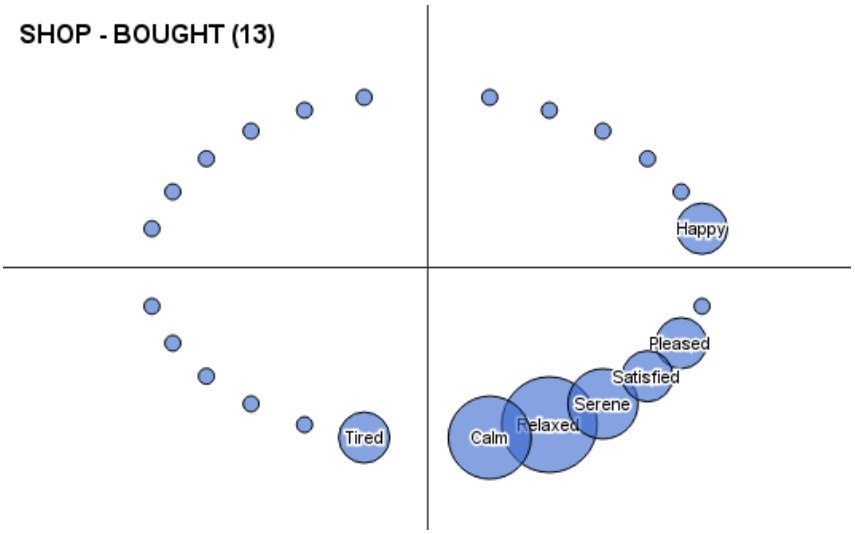


Figure 6.15: The figure shows felt emotions when players bought or sold something in a shop. Most emotions were recorded in the right-down sector indicating a passive-positive attitude for this outcome

with just 8% of negative emotions.

Random versus OpenStreetMap Algorithm

Due to the fact that the participants were split up in groups between two versions of spawning algorithm in the field trial, the categorization was also applied for the *Emotion Wheel* results. The *OpenStreetMap* version registered 79 log entries and the *Random* version 78. A differentiation between positive/negative or active/passive emotions has not revealed any major distinctions between these groups. In the *OSM* version slightly more positive and active emotions were recorded compared to the *Random* version.

Gamer vs Non-Gamer

Gamers and non-gamers were also separated to find differences in the target groups. 69 records in the gamer sector and 88 in the non-gamer group were logged. However, there was also no big emotional difference in this groups. 83% positive emotions were recorded for non-gamers and 72% in the gamer group.

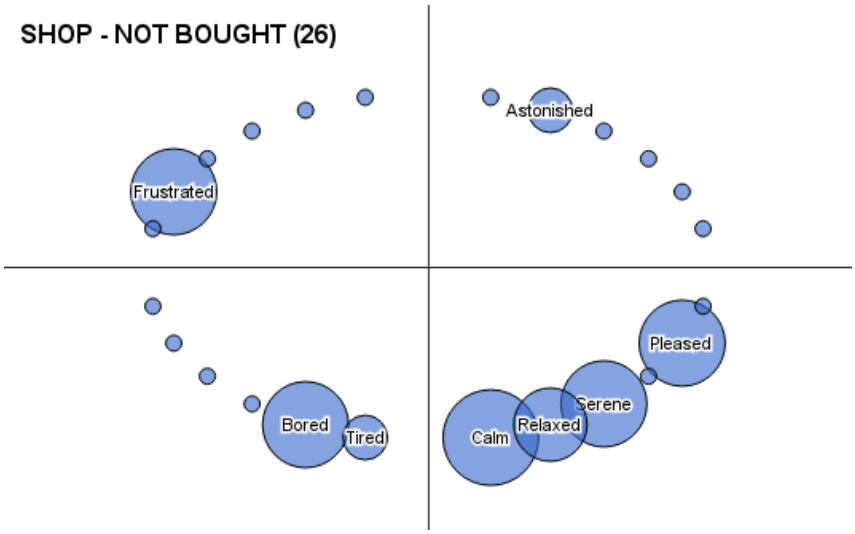


Figure 6.16: When players visited a shop and did not buy or sell anything, emotions in all quadrants were tracked with the majority in the passive-positive quadrant.

6.6 Interview

After every participant had finished the questionnaires, a semi-structured interview was held to receive additional qualitative feedback. The discussion was administrated by the supervisor of the field experiment and open questions were asked in the group to break the ice. During the discussion three questions were extracted from the necessities of the participants.

1. Have there been any problems while playing *Geo Heroes*?
2. What did you not like about *Geo Heroes*?
3. What did you like most about *Geo Heroes*?

Since the whole field trial was held in German also the answers were given in that language but translated for the reader. Answers for each question are described in a separate paragraph.

Have there been any problems while playing Geo Heroes?
Due to the fact that the game was still a prototype at the time of the

experiment, the participants reported several bugs and usability issues. One participant complained about the not updating distance to objects while walking. A head-up display (HUD) was displayed when clicking on the object where details were listed including the distance. The distance was not updated while walking and only another click updated the display with the actual distance. Other participants reported that the shop inventory was too large and they lost track. When opening a shop, things such as equipment could be bought. The problem here was that the equipment shop had separate items for each part of the body and additional ones for each level which were represented in a list view. Therefore, the list was quite long on the small mobile screen. Another issue was the status on the map screen, which was not shown to the user. Participants wanted to see the consumable status attributes not only on the battle screen but also on the map screen. This was reported to be important, since starting a battle without knowing the health points could lead to defeat. Especially non-gamers reported that getting started with the game was quite complex. Role-playing games are not easy to understand when played for the first time and a lot of functionality has to be scanned first to see the capabilities of the game. Most of the participants were unsure about the existence of resources and the usage of recipes. Everyone collected resources, but they were not sure about a profitable use case in the game. Also battle abilities were not clear for everyone, even though there was a dedicated menu entry to view all abilities with their effects. Another bug was mentioned with regard to the GPS. The GPS position froze after the screen had been turned off and woken up again. For a few seconds the movement of the character on the map screen was thereby stopped and participants had to wait until further interactions could be done. A missing indicator for the distance between player and object also led to problems when an impassable obstacle in the physical world prevented the interaction due to distance constraints. Another problem was noticed referring to the user interface. Opening the menu hid the map view and participants reported that a smaller menu would be more suitable to interact with the entries, while still having an overview of the objects around them.

What did you not like about Geo Heroes?
All design decisions were taken internally by developers before the field

experiment described in this thesis was done. Since a user centered design approach was chosen, the opinions of the participants play a central role in improving the game. Participants reported that an escape-from-battle button would improve their experience because the status values of enemies were not displayed and a battle defeat could often not be avoided. Another issue was related to the density of objects. Participants reported that enemies are too close to each other and a selection of two nearby enemies was not possible. The shop inventory was unknown and it was disappointing when nothing could be afforded. The continuous watching of the screen also reduced the game experience for participants while playing the game. Since headphones were used in the first half of the game-play session, participants concentrated more on the music and therefore on the game and reported collisions with other people in the city. Some participant were also feeling ashamed in the presence of others while watching the display and wearing earphones. Additionally, a few players remarked that they did not like the ambient music. Participants also claimed that equipment received from battle loot was not automatically equipped to a free equipment slot on the character. Especially gamers reported this kind of flaw, since they know this feature from standard role-playing games. Another feature participants were missing was the support of multiplayer mode. They all agreed that playing together and seeing each others on the map would increase their game experience. Participants also disliked the dialog with emotions for evaluation purposes, which popped up all the time after an interaction.

What did you like most about Geo Heroes?
This question was aimed at participants to report their favors in the game. Besides game content, also design and technical implementations were noted. For example, some participants reported the good choice of ambient music, which distracted them from everyday tasks in office. Others mentioned the high accuracy of the compass and GPS signal. The rotation and translation of the character was very smooth and therefore received positive feedback. The style of the map was also viewed positively, as it provided a better orientation on small ways without distracting participants with unnecessary map description and symbols. Of course participants also reported that they like this type of game-play, a combination of a traditional computer game and moving around in the physical world. They were engaged in doing

leisure activities such as eating ice cream while walking through the city and playing the game without the need of extra time. Additionally, players mentioned that they appreciate when games support physical activities such as walking in this case.

6.7 Results Summary

In this section significant and particular results are summarized for the sake of clarity for the reader. In the GEQ *Core Module* the components *Competence* and *Tension/Annoyance* showed significant results. There were significant differences between the versions (*OSM* vs. *Random*) and gamers vs. non-gamers concerning the *Competence* component. Participants felt significantly more competent with the *OSM* version than with the *Random* version. Gamers felt significantly more competent than non-gamers. The *t-test* for *Tension/Annoyance* showed that there is a significant difference in the gamer category. Gamers were much more annoyed than non-gamers. The *Post-game Module* of the GEQ showed only tendencies to significance in the version category in *Positive Experience* and *Tiredness*. *Positive Experience* was higher in the *OSM* version. *Tiredness* was higher in the *Random* version, and therefore the *OSM* version performed better with regard to player enjoyment. The PGFQ also had tendencies towards significance in the *Concentration* component. Non-gamers were more stimulated and their cognitive capacities were better utilized. *Player Skills* showed significance in the version category. Here the *Random* version performed better than the *OSM* version. A tendency to significance was measured in the gamer category. Non-gamers voted game interfaces and the fun factor in learning the mechanics much higher than gamers. The *Emotion Wheel* showed that the overall experience was more positive and passive than negative and active. Negative emotions were higher in case a player had lost a battle or no item was received from a resource deposit. In contrast, when players received the maximum amount of resources, they felt positive and also more active. Shop interactions also showed a big difference in emotions. When players had a successful transaction, 92% of them had positive emotions. A shop visit without a transaction lowered the positive emotions to 65%. The difference in the *Emotion Wheel* between the played version and gamer

category was extremely low. 82% of participants with the *OSM* version had positive emotions, whereas players of the *Random* version evaluated the game with 76% positive emotions. A slightly higher difference could be seen in the active and passive classification. The *OSM* version received 35% active emotions in contrast to the *Random* version, which only scored 28%. The gamers and non-gamers categories also showed a difference in the distribution of emotions. Non-gamers rated the game interactions with 83%, whereas only 72% of gamers had positive emotions. The group interview also showed some interesting aspects concerning usability and game design. Participants pointed out the things they liked and disliked in *Geo Heroes*. Additionally, participants revealed technical issues which were more or less linked to the received user experience.

7 Discussion

The results of the previous chapter are analyzed and discussed in the following sections. Since quantitative and qualitative methods were conducted, this thesis tries to bridge the gap between the results of each evaluation technique. First, the results of the conducted evaluation methods are interpreted. Afterwards, each single formulated hypothesis is being discussed in separate sections. At the end, the gained information is used to come up with concepts in order to improve *Geo Heroes* in the future.

7.1 Interpreting the Results

Evaluating *Geo Heroes* was done with different techniques to measure user experience and player emotions during game play. Although each method could be interpreted for its own, a combination in this thesis is preferred to prove evidence and to argue independently from gained results. On the one hand, evaluation was accomplished with the *Game Experience Questionnaire (GEQ)*, a tool for measuring user experience in traditional computer games. On the other hand, the *Pervasive Game Flow Questionnaire (PGFQ)* additionally provides the pervasive factor to consider the location-based attitude in *Geo Heroes*. Furthermore, the logged player emotions, which were tracked after each interaction, and the final interview with participants will also support the arguments for the interpreted results.

7.1.1 Immersion and Game Flow

Sensory and Imaginative Immersion from the GEQ was measured over all participants with a value of 1.67. The Immersion component in the PGFQ

achieved a value of 1.96 on a *Likert* scale between zero and four. Both components strongly relate to each other and do not show any significance or tendency in any category. In the GEQ, the component *Flow* reached a rather high value of 2.98, but also do not show any tendency to significance. Since the PGFQ uses components describing the *Game Flow*, it can be argued to calculate the average of all components to receive a comparable value to the GEQ's Flow. The average value of all PGFQ components achieved a score of 1.99. The results for immersion in the empirical study showed that GEQ and PGFQ provide similar results in all categories. Although the factor of GEQ is slightly lower than the one measured in PGFQ, it can be argued that both methods can be used to measure player immersion in a location-based game. Both average values are located close to the center in the *Likert* scale evaluation. Since immersion can be defined as the state of being in a virtual world receiving a continuous stream of stimuli (Witmer and Singer, 1998) and losing track of the physical world (Carrigy et al., 2010), the results of the final interview with participants help to give reasons for an average immersion value. Participants often reported that they almost had collisions with citizens during the game play session. Therefore, a shift between the virtual and physical world was necessary to prevent accidents. The two major problems were the smartphone displays, which had to be watched almost all the time to proceed in the game, and the ambient music when wearing headphones. Players were concentrated on the map to orientate themselves in the city and with additional sounds and music participants were not aware of the occurrences in the physical world. This forced participants to leave the virtual world when relevant things happened in the real world. A further problem which might have reduced the immersion was the disorientation of participants in unknown locations. They were forced to find the correct way in the physical world to reach the targeted virtual objects. The results of the GEQ in the post-game module emphasize the medium immersion value. In the post-game module the *Return to Reality* component was low which indicates that a shift back to the physical world was achieved without any problems. Games with a high immersion tend to let players remain for a while in the virtual world after they have stopped playing (Qin et al., 2009). Besides *Immersion*, also *Game Flow* was measured for *Geo Heroes*. From the results in the GEQ one can argue that player enjoyment was really high, due to the highly rated value of 2.98 in the *Flow* component. However, the results from the PGFQ

were not as high as the one presented in the GEQ. Since game flow can be seen as the optimal experience of player enjoyment independent from context and culture (Csikszentmihalyi, 1990) combined with appropriate aspects from user experience in computer games, a comparison between the *Flow* component in the GEQ and the overall result of components from the PGFQ should therefore be adequate. Unfortunately, a direct connection between these two values cannot be established. Calculating the average of all components in the PGFQ, a value of 1.99 is the result of the *Game Flow* factor in the game. Compared to the value of 2.98 from the GEQ, a significant difference is apparent for this factor. Despite the considerable gap of these numbers, it should also be considered that flow or game flow cannot directly be measured with a Likert Scale. Hence, flow represents the optimal experience of player enjoyment and therefore flow should be treated as a binary value, which means that a player is either in the flow or not, but never in between (Cox et al., 2012). This can also be argued with regards to immersion. While immersion progresses through game-play (Brown and Cairns, 2004), flow can only be achieved when immersion is at its peak (Jennett et al., 2008). Therefore, the given results for flow from the GEQ are not considered for the evaluation and the calculated result from the PGFQ is used, but only as an indicator until total player enjoyment is reached in the game.

7.1.2 Challenge

The challenge factor was measured in both questionnaires, GEQ and PGFQ. In the GEQ this component achieved an average score of 1.67. The *Challenge* component in the PGFQ reached an average value of 1.89. No significance was measured in the t-test analysis. Comparing these results reveals that there is only a slight difference between the overall average and the played version categorization. A bigger difference in the results is shown in the distribution of gamers and non-gamers. While gamers were more challenged in the GEQ during the field trail, the PGFQ shows that non-gamers were more confronted with challenges in the game. However, the problem with comparing the *Challenge* components are the views of the researchers. In the GEQ the questions in this component are asked in a more negative

tone, while positive and engaging questions were asked in the PGFQ. This would explain the different scores for gamers and non-gamers in both questionnaires. For example, in the PGFQ a question was ”The challenges in the game matched my skill level”, with a high value indicating a good user experience. In contrast and more negatively, the authors of the GEQ provide ”I thought the game was hard”, which only expresses that a high value leads to a negative experience and a low value indicates an appropriate level of challenge or a too low difficulty, depending on the participant's understanding. Another question from the GEQ was ”I felt time pressure”. This question can also be interpreted in two ways by participants. On the one hand one can argue that during the game tasks are given which have to be solved with time constraints and therefore provide a high challenge factor in reaching this goal in a given time. On the other hand, one can also understand the question in a way that time pressure emerges of the time limitations of the conducted field trial. Due to the nature of the constructed questions in the questionnaires, the ones asked in the GEQ are easier to understand but the PGFQ provides more application-oriented questions. The results in the PGFQ are therefore more detailed and descriptive for game designers as the ones in the GEQ.

7.1.3 Positive and Negative Experiences

While most of the questions in the GEQ relate to factors which belong to specific attitudes such as *Competence*, *Challenge* or *Immersion*, some of them belong to the positive and negative experiences during and after the game play. These components are named *Positive Affects* and *Negative Affects* in the GEQ core module and *Positive Experience* and *Negative Experience* in the GEQ post-game module. *Positive Affects* scored an average value of 2.36, whereas the *Negative Affects* were measured with a value of 1.50. After the game-play session, *Positive Experience* was measured with an average value of 1.31 and showed a tendency towards significance in the version category. The *OSM* version scored 1.63 and the *Random* version received a value of 0.92. The *Negative Experience* component did not show any significance and was measured with a value of 0.81. Those components measured in the core module describe the feelings participants encountered

during the game play session and will therefore be interpreted with the help of the log data represented by the *Emotion Wheel*. Since the log data also includes the emotions of participants during an in-game interaction, a comparison between the overall result of the *Emotion Wheel* and the core module results of the *Positive* and *Negative Affects* listed in the GEQ should therefore be possible. The 157 tracked emotions of all participants were split up in positive and negative emotions, where 78% were positive ones. This shows that in total the overall experiences were more positive than negative. In the GEQ the *Positive Affect* component scored an average value of 2.36, whereby the *Negative Affect* resulted in a value of 1.50. Comparing the results of the GEQ, it can be argued that there might be a relation between the values tracked in the log data and the values given in the questionnaire. However, it is hard to provide evidence for this empirical study. A further field trial with more negative than positive affects should be conducted to see if there is a connection between the results of the *Emotion Wheel* and the mentioned components in the questionnaires. Due to the fact that the log data point out the interactions during which more negative experiences were measured, one can adjust the settings in the game forcing these interactions to happen more often and reducing the amount of interactions where positive experiences were gained. For example, the designer can increase the difficulty of battles, with the result that a player loses more often. Afterwards, the results can be compared and discussed with the results presented in this thesis. After the game-play session the post-game module of the GEQ also provides questions for positive and negative experiences. In contrast to the core module, the post-game module assesses players' feelings after they stop playing. The results show that the average value over all participants is higher for the *Positive Experience* component than for the *Negative Experience* component. When splitting up participants in the categories gamer and version, the *Random* version is the only class where more negative than positive experiences were measured (Table 6.5). The *OpenStreetMap* version on the other hand, has the largest difference between the measured values of the positive and negative experience components. This result might point out the relevance of virtual content spawning in an appropriate environment in the physical world (Carrigy et al., 2010). However, although this is a promising result for future evaluation in this area, it should be noted that the evaluation was done with just a few participants. Furthermore, participants were not asked how they

felt before they started the game-play session. Such a questionnaire, which includes asking participants about the positive and negative experience before playing the game, could reveal important information on how the game version actually relates to the experience components and obtained results can thus be confirmed. There might also be relations between the positive and negative experience from the core and post-game module and the *Returning to Reality* component of the GEQ's post-game module. Since experiences measured in the post-game module only affects players' feelings after playing the game, the *Return to Reality* factor might have an impact on the disparity of results. The higher the measured value, the harder it is to return to the reality after game-play. Consequently, a high value in this component will lower the difference between the feelings in the core and post-game module due to the prolonged impact of the game after game-play has stopped. Unfortunately, this thesis can not provide evidence for this statement, but it is mentioned for reevaluation and future analysis in this area.

7.1.4 Object Interactions and Player Emotions

During the game-play session of the field experiment, participants logged their current emotions after each interaction in the game. Interactions were divided into battle, resource and shop outcome. As the results show, players experienced more positive emotions when a battle was won, which is unsurprising. However, the outcome was not only split up in winning or losing a battle. Additionally, a third outcome was introduced which is called BATTLE_WIN_SPECIAL and it describes the winning of a battle with special loot, such as potions or equipment, from enemies. The results revealed that between BATTLE_WIN and BATTLE_WIN_SPECIAL there is still a small difference in the emotions. While BATTLE_WIN resulted in 89% positive and 36% active emotions, the BATTLE_WIN_SPECIAL achieved 94% positive and 43% active emotions. This may look quite unimpressive because of only a small increase between these values, but considering the results from the resource outcome, where the maximum amount of resources has led to 58% logged active emotions, there might exist a relation on the axis of active and passive player emotions and special rewards. This

would mean that players are more actively involved in the game when special rewards are gained. As a result, such an active involvement of players increases motivation, which is relevant for continuing playing (Gee, 2003). It should be emphasized that even though losing a battle increases negative emotions of players, such outcome scenarios cannot simply be removed, since challenge is an important factor to receive game flow and thus an optimal gaming experience (Schell, 2008). Nevertheless, a game designer should be able to remove obstacles to gain a fun experience for users (IJsselsteijn et al., 2007) without neglecting important factors such as challenge. Possible examples to prevent a battle defeat in *Geo Heroes* would be to display the status of the enemy before starting the battle, so that players can weigh their odds. This feature enables players to better understand given scenarios and to learn from them, which can then increase the motivation to play the game (Gee, 2003). A further feature could be a flee-option with a given probability when battle has already started and there is no possibility to win against an overwhelming enemy. This would increase the chance to avoid negative emotions by protecing players against battle defeats. However, a flee-situation was not implemented and should therefore be evaluated in future work to see which emotions occur when successfully fleeing a battle. Collecting resources with different outcomes was also implemented in *Geo Heroes*. A player had a random chance to collect zero to three pieces of a resource from a deposit. As already mentioned above, when receiving the maximum amount of resources a player felt not only positive but also more actively involved. There was no big difference in the emotions as long as at least one piece was gained from a resource location. However, if a player unluckily received zero pieces, more negative than positive emotions were logged. 60% of emotions were negative, but still the remaining 40% belonged to a positive emotion. Participants were therefore asked why they had positive emotions even though they were not rewarded for their physical effort. As a surprising response, participants often did not care about the number of resources they gained. The reason for this is the same dialog which only differentiates from a successful collection by the displayed number in the dialog. This number was not explicitly highlighted as it was part of a sentence which had to fit in the small screen of a mobile device, which again could reflect sunlight, and was therefore often not recognized by players. Since sunlight has an impact on the visibility of text on displays, representative graphics should be used

for location-based games (Díaz et al., 2014) to avoid problems such as unsuccessful collections of resources. This should also be considered when a player receives special rewards after winning a battle, since it was not always recognized by participants as it was only displayed in form of a text next to some general information. Other interactive game objects were shops. In a shop, players could buy items and in return they had to pay the specified amount of gold, which was gained by defeating enemies or selling items. The interaction was divided into two outcomes. Either players could interact with the shop to buy or sell something, or they could leave withou doing so. When players did not buy items, it indicated that they either did not have enough gold to buy something or the items were not of interest. Out of the 39 shop interactions, 13 transactions were logged. In this case the results were very clear. 92% of the emotions were positive and also passive. In contrast, when nothing was bought or sold only 65% were positive, but still 81% was passive. When looking more closely on the negative side of the *Emotion Wheel*, it is clearly evident that the emotions *Frustrated* and *Bored* were almost always chosen from participants. In the group interview participants reported that shop inventory could only be viewed when being in close vicinity to a shop. However, the game only provided one type of shop per square kilometer. Hence, participants had to visit the shop explicitly to look at the shop's inventory to buy or sell items. Unfortunately, prices were higher than the amount of gold of participants and they therefore had to leave without a transaction. This was frustrating and discouraging for some people and they reported in the discussion that they had to travel to a distant location without gaining an appropriate reward (Procyk and Neustaedter, 2014). Others, who felt bored, reported the long item list from the shop's inventory. Since each equipment part has levels from one to ten, the *Equipment Store* provided a lot of items to buy. The long list representation on the small mobile screen was unacceptable for participants and they got bored when trying to find something useful for their current character level and amount of gold. In order to avoid such scenarios, game designers could give information on the shop's inventory even when not being within interaction distance. Furthermore, shops should only list items which are appropriate to the player. For example, a character level one should not be confronted with items for a character level ten. Also a grid view with the most important information such as price, icon and level would be better for an item overview instead of using a list view.

7.2 Overlapping the Physical and Virtual World

One primary goal of this thesis is the evaluation of the spawning algorithm. For this purpose, participants were divided into two groups. One group played with the *OpenStreetMap(OSM)* version, the other with a *Random* version. The *OSM* version provided an algorithm which is responsible for spawning virtual content close to an appropriate physical object. The *Random* version, on the other hand, only used a spawning algorithm for content creation based on random probability. The *OSM* version had the goal to achieve a better user experience, since locations should be in relation to the content of the virtual world (Carrigy et al., 2010). In all evaluation methods, the version diversification was included in the field experiment and statistical significance was measured where it was suitable. It can be seen from the results in the core module of the GEQ that there are some differences between the components, but only the *Competence* component showed a significant difference between the groups. The average value of *Competence* was 1.64, but the value in the *OSM* version scored 2.00. In contrast, the *Random* version only scored 1.20. Other components in this module also had a tendency to perform better in the *OSM* version, including the *Challenge* component, where a lower value is preferable for a good user experience. In Figure 7.1 an overview of the different results in the components of the played versions is presented. In the post-game module of the GEQ there was a tendency towards significance in the *Positive Experience* and *Tiredness* component. The *OSM* version scored 1.63 and the *Random* version achieved only a value of 0.92. *Negative Experience* was encountered to be higher in the *Random* version, but the t-test does not show any tendency of significance. In Figure 7.2 the results and differences between the played versions are shown. In contrast to the GEQ, the PGFQ provides results which are totally mixed in this classification. Only the *Player Skills* component showed a statistical significance, but the *Random* version received a higher score (2.50) than the *OSM* version (1.94). In Table 6.8 the results are summarized for the components in the PGFQ. A conclusion on a component level would not be reasonable. Hence, a per question analysis is being considered to uncover some relevant information. Although there are several questions in the GEQ with a significant difference, most of them will not be used for extracting detailed information, since these questions are

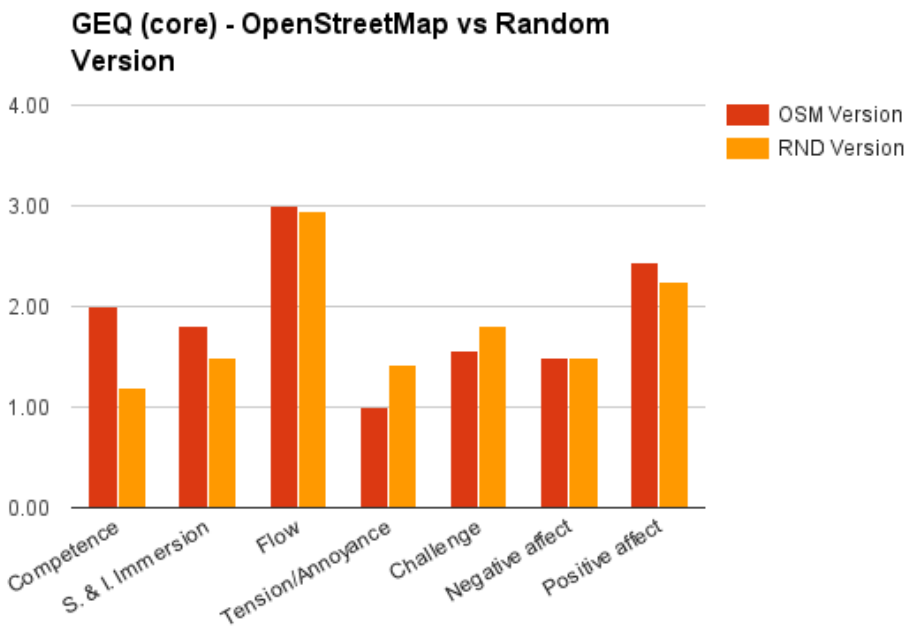


Figure 7.1: A comparison between components of the Core module in the Game Experience Questionnaire in relation to their content placement algorithm (OSM = OpenStreetMap; RND = Random) presented in a bar chart. The components for the OSM achieved better results for player enjoyment. A high value in the *Challenge* component and in the *Tension/Annoyance* component means a low user experience.

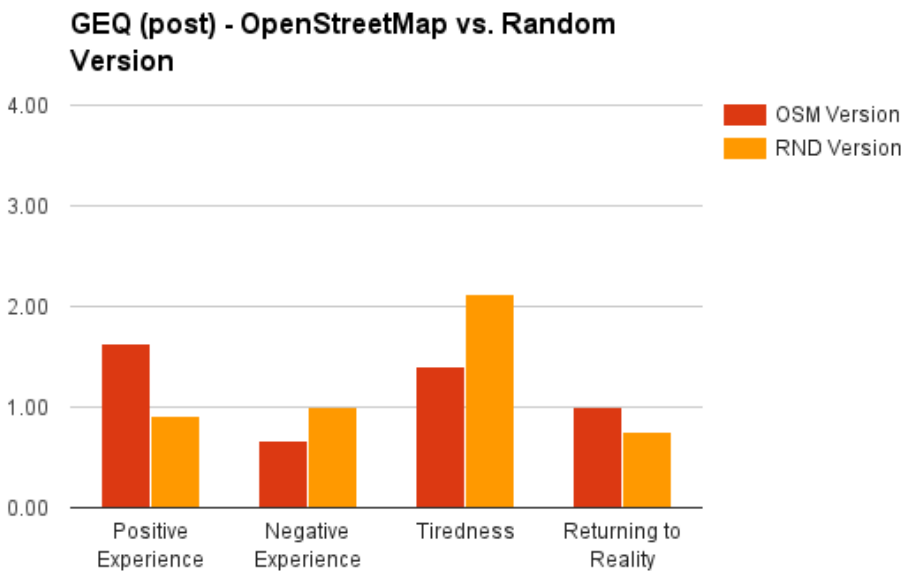


Figure 7.2: A comparison between components of the post-game module in the Game Experience Questionnaire in relation to their content placement algorithm (OSM = OpenStreetMap; RND = Random) presented in a bar chart. The components for the OSM achieved better results for player enjoyment. A high value in the *Negative Experience* component and in the *Tiredness* component means a low user experience.

described rather too simply for such an analysis. The only question which provides useful information in this area is the question:

> ***Q:*** *I had the feeling that I had returned from a journey.*

This question received a value of 2.4 for the *OSM* version and only 1.00 for the *Random* version. Statistical significance is also apparent. A journey can be seen as an adventure adapted to a coherent environment, which is given due to the provided algorithm based on map information. In addition, the questions in the PGFQ are conducted to reveal some relevant notes, which could prove the usefulness of a spawning algorithm based on map information. From the 34 questions in the PGFQ only three achieved significant difference (< 0.05) in the t-test of the played game versions. Two of them also showed significance or a tendency towards it in the gamer category but are not considered for a further analysis due to the question's nature. Therefore, one question is left, which is related to the *Immersion* component:

> ***Q:*** *The game enabled me to shift focus between the virtual and physical parts of the game world without losing too much of the feeling of immersion.*

This question was rated with 2.80 by participants who played the *OSM* version. In contrast, players of the *Random* version reached a value of 1.00 on average. This could be the proof that users tend to play a game which provides game content at appropriate locations in the real world. The log data was also analyzed to discover some significant differences. But since the emotions were tracked as a response to direct interactions, no relevant information can be extracted in this case. Both versions achieved similar results in *positive/negative* and *active/passive* emotions. However, further research needs to be done in future to examine the results of this field experiment. As a suggestion, more participants should be recruited to reduce the chance of a coincidental result and additionally a qualitative evaluation method should be used to confirm the findings.

7.3 Gamer and Non-Gamer in Pervasive Games

Although a game was evaluated in this thesis, also non-gamers were invited to the field experiment. Due to the fact that location-based apps or games differ from traditional computer games, a classification between gamers and non-gamers for measuring user experience could be of interest. The greatest difference between these groups is the experience factor in gaming. *Geo Heroes* is a typical role-playing game in a medieval-fantasy genre and players who are familiar with this type of game can more easily adapt the game mechanics than non-gamers. Despite that, gamers play their games usually in an indoor environment, but are forced to leave their comfort zone to play *Geo Heroes*. In contrast, non-gamers do not play games at all or quit gaming as they grew up. They often report their limited leisure time and that they do not have time to play traditional computer games. *Geo Heroes* was developed with the advantage to standard computer games, that they can be played outside during everyday tasks. Therefore, also non-gamers were invited to this field experiment. In the core module of the GEQ the t-test provided statistical significance for the *Competence* and *Tension/Annoyance* component, where gamers reached a higher value in both. *Competence* achieved a value of 2.05 in the gamer category. Non-gamers reached a score of only 1.32 in this component. Reviewing the questions of the *Competence* component shows that gamers felt more competent with the game than non-gamers. However, gamers were also more annoyed than non-gamers. In the group interview, especially gamers reported problems related to user interface and game mechanics as they compared the still developing game *Geo Heroes* with already released traditional games. Non-gamers, on the other hand, did not report that amount of problems and reached only an average score of 0.67 in the *Tension/Annoyance* component. In contrast, gamers reported an average value of 1.83. The post-game module of the GEQ did not show any statistical significance on the component level for the gamer category. There are only slightly higher measured values in the components *Positive Experience*, *Negative Experience* and *Tiredness* for gamers. An increased number of participants would maybe show a tendency on how gamers and non-gamers feel after playing a location-based game and hence, statistical significance could be revealed in one of these components. The PGFQ showed significantly higher values in the *Concentration* ($G_{yes} = 1.63; G_{no} = 2.33$) and *Player Skills*

($G_{yes} = 1.82; G_{no} = 2.49$) component for non-gamers. It should also be mentioned that results from each component were higher for non-gamers than for gamers. Additionally, the *Emotion Wheel* showed that there were more positive emotions in the non-gamer category. This could be an indication that *Geo Heroes* and location-based games in general are more popular for non-gamers than for gamers. Unfortunately this cannot be proven as long as participants are recruited in a field experiment. Such an assertion can better be proven when people are voluntarily playing the released version of the game for a longer period. Therefore, the evaluation in this thesis cannot conclude whether gamers or non-gamers are more likely to play *Geo Heroes*. Nonetheless, the results in this thesis show that both gamers and non-gamers achieved similar results in most of the components. More concrete information on how to improve the game was contributed by gamers, since they are more critical due to their experience. In contrast, non-gamers also provided useful results, since they are not biased from traditional computer games and are a potential target group for location-based games.

7.4 Quantitative Evaluation Methods for Pervasive Games

In the past, pervasive games were more highly esteemed in the area of research. Developing such games is not an easy task, since there are a lot of hardware and software issues depending on the game design and there is no 'out-of-the-box' framework or game engine for this kind of games. Even if designers can manage the requirements of such a game, there are still further complex tasks which have to be solved such as testing and evaluating. However, pervasive apps or games are getting more and more important and Jegers (2007) predicts that these will become an everyday phenomenon in future. For example, Pokemon Go[1] was released in early summer of 2016 and players have crashed servers due to the immense demand. Despite the great success of Pokemon Go, critique and user reviews are not always positive[2], even though designers conducted a long beta test period before.

[1] http://www.pokemongo.com/

[2] http://www.metacritic.com/game/ios/pokemon-go

A generally established evaluation method for such games could prevent bad critique by tackling design problems already during development.
Geo Heroes was therefore evaluated with a model specially designed for pervasive games presented by Jegers (2007), which is based on the game flow concept from Witmer and Singer (1998). This model was then transformed into a questionnaire with 34 questions called PGFQ. In addition, *Geo Heroes* was also evaluated with the GEQ from IJsselsteijn et al. (2007), which is usually used for traditional computer games. As shown in the result chapter of this thesis, both methods were promising in detecting components which should be improved to raise game experience or player enjoyment. However, there were some issues detecting concrete design problems with the standard GEQ. Questions such as

Q: *I felt happy.*

can of course measure the positive aspects in a game, but if such questions are rated low, a further analysis for detecting problems is not easy. The advantages of the GEQ are the simple questions and the quick assessment during or after a game-play session. On the other hand, GEQ may lack potential characteristics of a pervasive game and therefore the PGFQ was used to fully assess a player's experience when playing *Geo Heroes*. Besides the additional aspects the PGFQ is covering, questions are described in more detail and are therefore easier to interpret for game designers if a low value is scored. For example, the question

Q: *The game provides appropriate rewards for my effort.*

can draw conclusions about the rewards gained after an object interaction and the distance the player had to cover for it. While such questions give designers a better understanding of a player's needs, detailed questions also result in a longer duration to fill in the complete questionnaire. The bigger the time frame until each question can be answered, the more problems can occur in detecting the felt game experience of participants during the game-play session (Frommel et al., 2015). Nonetheless, the PGFQ provided proper results for the field experiment. These findings were confirmed by the group discussion and the results from the emotion wheel. Hence, it can be concluded that the PGFQ can be an important tool to provide a good assessment of the measured player enjoyment of *Geo Heroes*. Additionally,

such a questionnaire can improve the game by detecting concrete design issues.

7.5 Support Quantitative with Qualitative Methods

Quantitative evaluation methods, such as questionnaires, are often used to measure user experience in standard software applications. The goal of these questionnaires is to measure traditional 'productivity' metrics (IJsselsteijn et al., 2007). Although such metrics were successfully applied to the gaming genre, further factors are relevant to measure user experience or enjoyment in games. Additionally, there is a continuously growing competitive market in the gaming industry. These facts increase the attractiveness of finding new methods and models for companies and researchers in this area. However, the market for location-based games is still in the early stages of development. Furthermore, the technical requirements of such games are often not easy to overcome. Thus, there is no great variety of methods to evaluate location-based games. In this thesis, both quantitative and qualitative methods were applied to measure user experience in the location-based game *Geo Heroes*. Questionnaires such as the GEQ and PGFQ were conducted to assess participants for a quantitative analysis, whereas emotional responses and a group discussion provided information for a qualitative study. As already described in the previous section, GEQ and PGFQ can be useful tools for a quick assessment of player enjoyment. Nonetheless, such questionnaires fail to provide enough information for game designers. Questions in questionnaires such as the GEQ are often formulated in a too general way and a direct assessment of concrete design or usability issues can not be determined. Even when questions are formulated in more detail, such as in the PGFQ, some relevant points on how to improve the next version of the game can remain open. If these issues cannot be tackled, at-game frustration such as struggling with the user-interface may occur (Gilleade and Dix, 2004)(IJsselsteijn et al., 2007), which will lead to a bad user experience. Furthermore, questionnaires can interrupt the experience of flow and gained results can be deviated (Frommel et al., 2015). Moreover, research has shown

that questionnaires do not always provide useful results when measuring factors such as immersion or presence (Huhtala et al., 2012)(McCall et al., 2004). In order to tackle these problems, qualitative methods were introduced to evaluate user experience in *Geo Heroes*. Capturing player response was achieved through the usage of an in-game dialog presenting various emotions. These dialogs appeared after an interaction with a game object such as battling with an enemy or collecting resources. The chosen emotion was then categorized to the type of interaction including the outcome, for example a battle which was won. Of course, a post-game questionnaire could ask for players' feelings and emotions after such an interaction, but the variety of different interaction outcomes and time elapsed between game-play session and questionnaire would probably bias the results. Furthermore, integrating such questions would maybe transform the qualitative method to a quantitative one. The goal of using such an in-game reporting tool was to directly assess players' emotions when they were interacting with an object, and thus a truly felt emotion could be captured. Additionally, the question asking for the currently felt emotion does not pop up just once, as a questionnaire would implement it. Instead, several emotions can be tracked and a pattern might be discovered. Possible patterns would be for example, emotions over time or emotions at specific locations. This could additionally draw conclusions about concrete problems in game design. In the evaluation of *Geo Heroes*, these patterns were not elaborated in this first field experiment due to the short playtime and the different locations which were visited by participants. However, such additional information might be useful as it reveals potential problems in the game and should therefore be considered in future evaluations. In *Geo Heroes* the *Emotion Wheel* helped to confirm results from the conducted questionnaires. For example the non-gamer results in the components of the PGFQ were higher than for gamers, which was also evident from logged emotions in this category. Beside the usage of tracked emotions as a qualitative method, a group interview was additionally conducted after the game-play and questionnaire session. The interview was held informally and every participant was free to speak. Starting such a conversation is always a problem as people do not know where to start or what to say. The supervisor therefore asked some starting questions such as "What did you not like about Geo Heroes?" or "Has there been any problems while playing?". Such questions funtioned as an ice-breaker and helped participants to say their opinions. Naturally,

people agreed when someone noted a problem and also provided additional information. However, such qualitative feedback was important for *Geo Heroes* as it revealed further unknown issues. Also, the results from the conducted questionnaires were often confirmed by the given answers from the group discussion. Moreover, some answers helped to understand some low ratings in the component analysis from questionnaires. It can therefore be concluded that interviews are an essential tool to measure positive and negative experiences. Although such interviews are often conducted in a one-on-one conversation, *Geo Heroes* profited from the many answers in the group discussion. Even if there is a risk that results can be biased from more extrovert participants, a group interview has also some advantages for location-based games. First of all, every participant can do the field trial at the same time. Otherwise, people have to be invited over a bigger time interval and the evaluation is unnecessarily lengthened for supervisors and other responsible people. Second, the time delay is shortened between game-play session and discussion which helps participants to remember more relevant information during playtime, since the possibility of interruptions can therefore be reduced (Frommel et al., 2015). Third, the overall waiting time is minimized for participants and chances are higher to recruit more of them. Fourth and last, when playing a location-based game especially in public open spaces, quiet waiting rooms or rooms for a one-on-one discussion are often not available and therefore a group interview in a nearby venue is easier to conduct. However, the more evaluation methods are conducted, the more participants are challenged in their concentration. For example, the results of the *Emotion Wheel* provided valuable information for the game improvement, but also annoyed some players and maybe spoiled their experience in flow or immersion. Nonetheless, it can be argued that qualitative methods strongly support the usage of quantitative methods. Especially in the early stages of development, developers and game designers can receive a lot of information in design or technically related issues.

7.6 Improving Geo Heroes

The primary goal of this thesis is to improve the location-based game *Geo Heroes*. After analyzing quantitative and qualitative evaluation methods,

it should be possible to conclude design decisions for the next iteration in a user-centered design process. Since the PGFQ has proven to be the most expressive instrument in discovering potential problems in player enjoyment, the component structure of this method is used to arrange concrete solutions to their appropriate component or components.

Concentration

In order to provide tasks that feel more important, the game should include a meaningful story line. A story could be split up in various sub tasks which are of importance for the overall game. According to the aspect of proving stimuli from different sources, physical objects, which can be found on the map, should be introduced and aligned with the game content. This approach is also used in *Geocaching*, where real players seek physical objects placed by other players. Additionally, *Geo Heroes* should make usage of audio GPS, where a voice prompts the player to the desired location. This would reduce watching the mobile screen all the time and hence support switching between in-game tasks and surrounding factors in the game.

Challenge

The game should provide difficulty settings for inexperienced and experienced players. This could increase the game experience for both non-gamers and gamers. Each player can then adjust the difficulty to their personal skills. Hence, the challenge factor would be at an appropriate level for every player. If player skills increase over time, the difficulty level can also be increased in the game settings. In order to prevent losing a battle due to unknown status attributes of enemies, players should get information about these values at any time. A head-up display for status attributes should be implemented and shown on the screen when a player selects an enemy. Players should also choose their path depending on their interactions. A player's choices, for example in dialogs, should have impact on the character's attitudes and progress. This could support players in the creation of own game scenarios and pacing. Furthermore, the game should provide the availability of user-generated content. Players would then be able to create their own stories in their city, which is probably a better approach than using just an algorithm based on map information. Since inhabitants have

a good knowledge about their surroundings, they can create scenarios for other players which fit to the underlying environment. For example, content creators could tell the history of their city.

Player Skill

In games, players have different skills or learn them differently fast, but it is nevertheless important to satisfy everyone who is playing *Geo Heroes*. An optional in-game tutorial teaching relevant game interactions could be a good starting point to create an equal base for all players. Through game progress, player skill should be improved and battles should require knowledge of enemies and items. This knowledge could be increased by finding relevant information of game objects in the game world. A further improvement in skill can be the implementation of more complex battle scenarios. For example, a two versus two battle, where a player has companions and tactic is required to defeat the group of enemies. The effort of players should also be rewarded appropriately. Therefore the player should be aware of a shop's inventory before traveling a long distance. Also, when collecting resources, a player should receive at least one piece from a resource deposit. Game interfaces and mechanics should also be improved especially for the requirements of gamers. As an example, there should be hot-keys to immediately switch to a specific menu. Reducing the size of buttons could also establish a better overview of the overall user interface.

Control

Improving the control factor in a game can be achieved in various ways. First, character customization can give players the ability to fully control the appearance of their character. Hence, in location-based games it is important that the movement of the character works fluently. Since the player's movement is crucial for the positioning of the avatar, GPS improvements can lead to a better user experience. For example, the minor bug with the GPS re-positioning after the screen turns off could improve the control factor. Player actions should also have an impact on the game world. Therefore the player should be able to create and add game content, such as buildings, which persist in the overall game. A picture of the current status in the game world can also be provided by implementing an overview map or enable a

higher map zoom including low poly map overlays to prevent performance problems.

Clear Goals

The only goal the players had in the field experiment was to reach level five with their character. This goal was not clear enough as it does not describe the reason why a player should do that. In order to provide clear goals, a story or quest series could enhance this factor. Hence, players are aware why they are playing this game and what they have to do to fulfill these goals. Also achievements, which represent personal intermediate goals, should be available and unlocked if succeeded. In a multiplayer scenario these can then be shared across other players to gain status in the community. This is also used in a wide range of game platforms such as STEAM[3]. *Geocaching* is using a similar approach for their players.

Feedback

The *Feedback* component from the initial *Game Flow* approach by Sweetser and Wyeth (2005) was not further elaborated in Jegers (2007). Nevertheless, it is important to provide continuous feedback for players when required. In *Geo Heroes*, the status attributes of players and other combatants should be presented to the player when needed. Additionally, during a turn-based battle, players should be provided with a log of battle events to better understand the current situation or why they have lost.

Immersion

Immersion is essential when playing games. In order to enhance immersion, players should be emotionally involved in the game. A story line using real world artifacts which are based on historical or cultural events could establish a deeper immersion. It is also important that players' actions should not result in a violation of social norms. A typical example from the evaluation was the screen watching issue while wearing earphones all the time. Participants felt embarrassed while playing the game and solutions should be implemented to prevent this kind of feelings. Context switching

[3]http://store.steampowered.com/

between real world and game world actions should be easy and should not reduce immersion. This can be provided by a minimal use of smartphones and of course an appropriate game content which overlaps with the physical world.

Social Interaction

Even though the current implementation of *Geo Heroes* does not provide social interactions due to the lack of multiplayer support, it is a relevant feature for a successful location-based game. Participants in the evaluation were also expecting to see a version of *Geo Heroes* which supports multiplayer functionality. However, it is not enough to have players communicate via a chat or to just make them visible on a displayed map.

> *"A pervasive game should support and enable possibilities for game-oriented, meaningful and purposeful social interaction within the gaming system."* - Jegers (2007)

It is therefore relevant to implement such social interactions. A possible form of interaction is trading with other players. This would enable a cooperative game-play in *Geo Heroes*, including factions or guilds. In order to provide more competitive social interactions, player versus player battles, also known as PvP, should be implemented to keep the motivation of players high. Moreover, the PvP approach can be extended to city-wide tournaments with a knockout system.

7.7 Discussion Summary

In order to answer the established hypotheses from section 3.2, the discussed results are summarized and referenced to their related hypothesis. It can be confirmed, that suitable quantitative methods, such as questionnaires, provide a good estimation to measure user experience for *Geo Heroes* **(H1)**. The *Game Experience Questionnaire (GEQ)* and *Pervasive Game Flow Questionnaire (PGFQ)* both show similar results in their components, even though some of them are differently interpreted. In the *Challenge* component for example a high value in the PGFQ means a better user experience, whereas

in the GEQ a high value stands for a low user experience. Nevertheless, it is obvious that quantitative methods are easy and useful tools for a rough estimation of user experience in gaming. The more participants a game development team can hire, the more precise the outcome will be. Also questions regarding a classification of users can be integrated in order to satisfy requirements of different groups. By providing appropriate tools such as a personal interview with participants or the *Emotion Wheel* for assessing players' emotions, it can be confirmed that suitable qualitative methods indicate whether players like or dislike the game in specific situations or at all **(H2)**. Due to the usage of the *Emotion Wheel* it was easy to gain a lot of qualitative feedback, for example which object interaction and outcome is responsible for positive and negative involvement in the game. Additionally, the group interview revealed some technical and design issues which influenced the game experience while playing *Geo Heroes*. As already discussed before, the *Pervasive Game Flow Questionnaire* proved to be the more appropriate tool to measure user experience in pervasive games **(H3)**. Even if the content is more complicated for a user to answer, it provides much more feedback than the standard GEQ. Designers can also benefit from single answers in each question, since they often describe concrete issues of a game design element. Furthermore, the PGFQ supports questions regarding the context switch between the physical and virtual world, which is a major attribute of pervasive games and gained information can be very valuable for researchers and game designers. Beside the results of quantitative methods, it can be said that qualitative methods have supported the answers and statistics of conducted questionnaires **(H4)**. This can be seen especially in the group interview. Relevant information is often lost due to the constrained possibility of answering in questionnaires. The interview helped to assign and confirm the already discussed problems to questions and components of the questionnaires. The hypothesis which describes the increase of overall user experience with the algorithm based on map information can not fully be answered **(H5)**. Of course, results indicate more positive emotions in the *Emotion Wheel* in contrast to an algorithm based on coincidence. Also, the GEQ provided better results for the *OSM* algorithm than the *Random* one. However, in most cases there were no significant differences and the group of participants was really small. Gained results also depend on the control group and the provided algorithm in the control group. Maybe a parametrized implementation of the *Random* algorithm could increase the

user experience and immersion. The differences in the *Emotion Wheel* were also not significant enough to draw a conclusion. Nevertheless it can be said, that appropriate virtual content at appropriate physical places can raise the immersion. Switching tasks from physical and virtual world can be easier in such scenarios and overall game experience can be increased, but it strongly depends on the nature of integration of physical places or objects. The last hypothesis, which describes a better user experience and an easier access for players who are familiar with traditional games can only partially be confirmed **(H6)**. It can be said that experienced players have easier access to a role-playing game since they know the game mechanics and most of the features which were provided in *Geo Heroes*. However, the user experience and hence player enjoyment was lower for gamers than for non-gamers in the *Emotion Wheel* and in the PGFQ. Also, the *Tension/Annoyance* component showed a significant difference and gamers had a much higher value than non-gamers. Gamers are more critical to a game in a prototype stadium, since they compare the game with already released ones. Non-gamers, in contrast, are more excited about game mechanics, and learning the game is much more fun for this group. Therefore, non-gamers often reached higher user experience values in the questionnaires.

8 Conclusion

The aim of this thesis was to find out how a pervasive role-playing game named *Geo Heroes*, which provided a content placing algorithm based on map information, can be improved in a user-centered design approach. Literature was reviewed to examine related work in the area of pervasive and thus location-based games, but also in the area of traditional computer games. After analyzing various approaches, hypotheses were formulated and the most applicable approaches were selected for an empirical study with participants to measure user experience while playing *Geo Heroes*. A questionnaire for traditional games (IJsselsteijn et al., 2007) was used after the game-play sessions took place. Subsequently, a questionnaire for pervasive games was provided, which was created out of the given criteria and components described in Jegers (2007). Beside these quantitative methods for measuring user experience, also qualitative methods such as tracking of emotions (Baillie et al., 2011) and a group interview were conducted to gain detailed information about felt player enjoyment during playtime. Participants were divided into categories to see whether there are significant differences between those. One category divided participants into gamers and non-gamers, the other one divided them according to the version they played. Either they played the experimental version with the algorithm based on map information or the version with a random algorithm for placing game content.
Results show that a questionnaire for traditional computer games is partially useful, even if it is simpler and faster to fill in. In contrast, a questionnaire for pervasive games, where questions are described in more detail can achieve results which are more useful for a user-centered design approach. However, the more detailed a questionnaire is, the more time and concentration is required for participants which could lead to problems remembering emotions such as immersion during playtime. Hence, direct assessment of

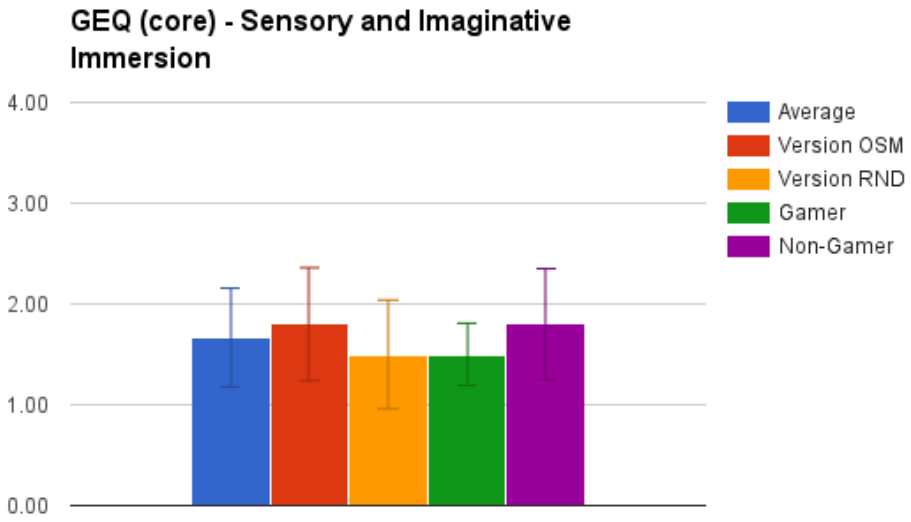
player emotions during the game-play session has shown to be an appropriate tool for a precise evaluation of positive and negative experiences after in-game interactions. Since not every need can be covered in standardized questions or limited emotions, a group interview was applied and revealed additional potential problems which should be tackled for the next step in design and implementation of *Geo Heroes*. All in all, it can be concluded that questionnaires for traditional and pervasive games are helpful tools to measure user experience in location-based games, but it has to be added that pervasive questionnaires are more appropriate for an evaluation. Captured emotions and interviews are useful instruments to assess player enjoyment during playtime. Furthermore, qualitative methods provided evidence of the gained results of the quantitative methods. A location-based game which uses map information is not significantly better than one with an algorithm using random places for virtual content. Despite that, it can be argued that immersion can be kept higher when virtual content is placed at appropriate physical locations, since the switching between the physical and virtual world is easier. In the category of gamers and non-gamers, there is also no significant difference in player enjoyment between the groups. Nevertheless, the results show that gamers are more critical due to their experience and are therefore a relevant group to gain deeper insights in design problems. The problems of non-gamers, who are also a potential target group of location-based games, should also be considered. Even if they do not detect problems in a fine tuning process, they can point out significant problems in general topics such as the lack of explanations or a slow learning curve.

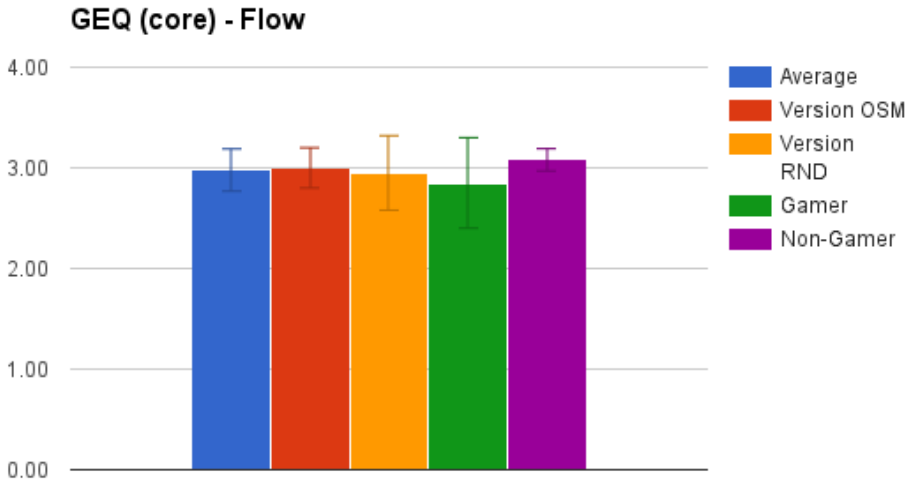
Although the formulated hypotheses are answered, there are a few more things to mention for a possible future work and suggestions for other researchers conducting a field experiment in this area. First, it is hard to find participants for playing a location-based prototype. Social media, newsgroups and landing pages were used, but playing such a game requires extreme effort. Arranging that everyone is at a defined place at the same time is hard, especially when participants are working in full-time jobs. It is therefore necessary to assess participants quickly with small effort. Even though Moser et al. (2012) did not evaluate location-based games, a rapid assessment should be designed for frequent field experiments, such as log data from Baillie et al. (2011), where results can be quickly evaluated. A possible solution could be the placement of the mobile app in the app store

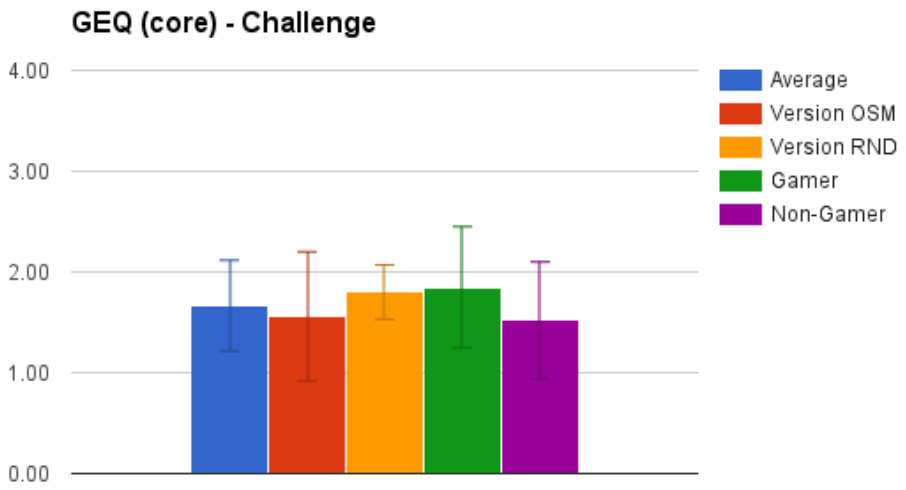
as an ALPHA or BETA version, and players have to agree to the usage of log data and further analysis. Although it could be problematic if such a field trial is not supervised, the amount of time hiring participants could be reduced to a minimum. Second, even when trying to create a new game that is promising to be a better version of a well established game such as *Geocaching* with core elements such as competition as described in Lund et al. (2010), it is not guaranteed that players will tend to play such games voluntarily in their leisure time. Therefore, results which are gained from recruited participants will probably not be the same as from players who are playing the game voluntarily. It may be obvious, but location-based games such as *Geo Heroes* can bring participants to their physical limits. It is suggested to provide drinks and food after the game-play session to keep them in a good mood, otherwise results can be biased when participants feel unwell. Another factor which may influence the results are weather constraints and they should be integrated if field evaluations are conducted more often.

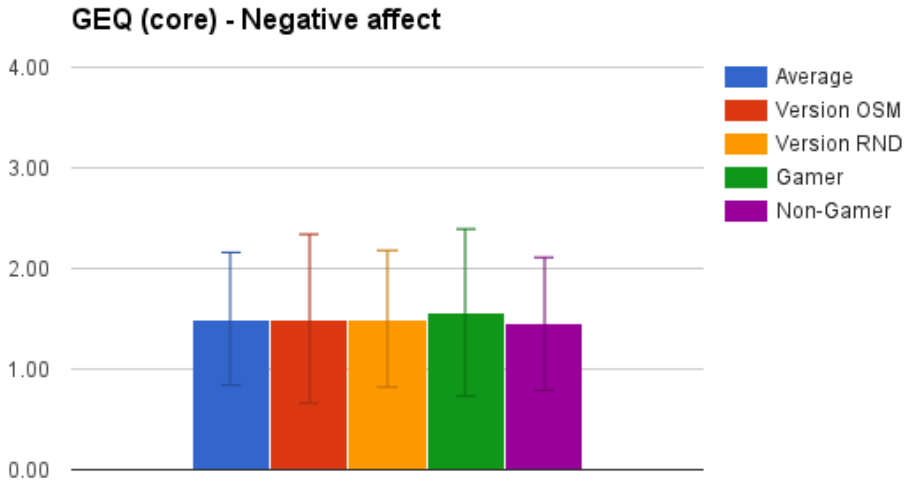
Taking all together, it can be concluded that existing tools for measuring user experience or player enjoyment, as it is sometimes described in literature, provide useful results for further improvements of a location-based game. This thesis has shown the potential of a user-centered design approach for a game which is still under development. Many issues were detected which would not have been uncovered without such a technique. The future of location-based games such as *Geo Heroes* therefore strongly depends on the opinions of users which should be frequently assessed to reduce the effort in design and development.

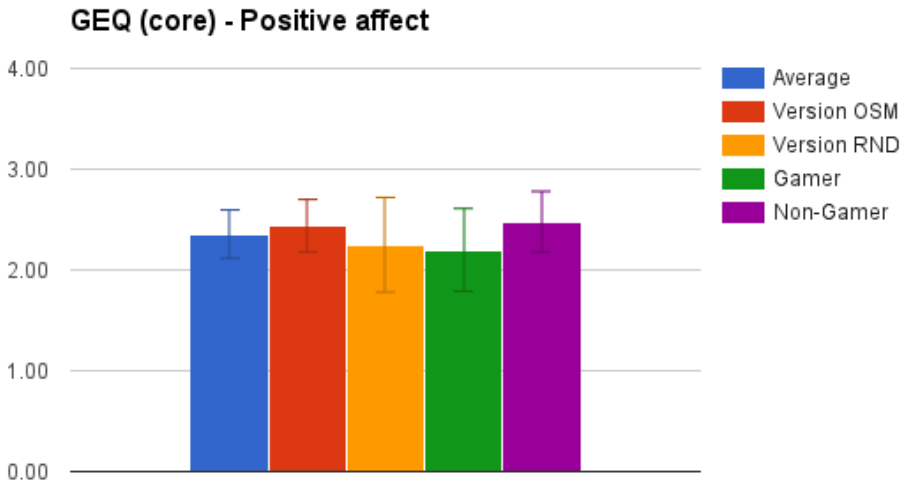
Appendix

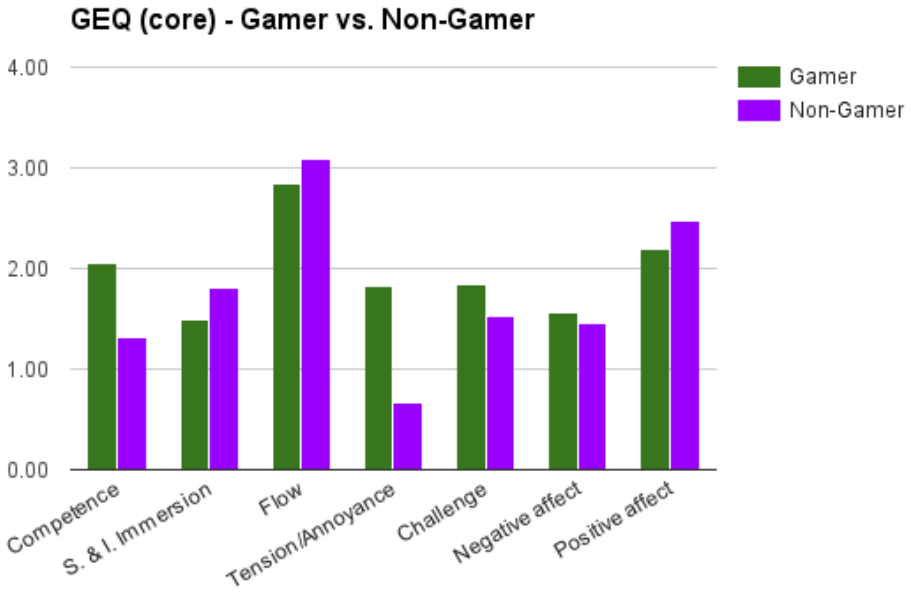

GEQ (core) - Sensory and Imaginative Immersion
4.00
3.00
2.00
1.00
0.00
Average
Version OSM
Version RND
Gamer
Non-Gamer

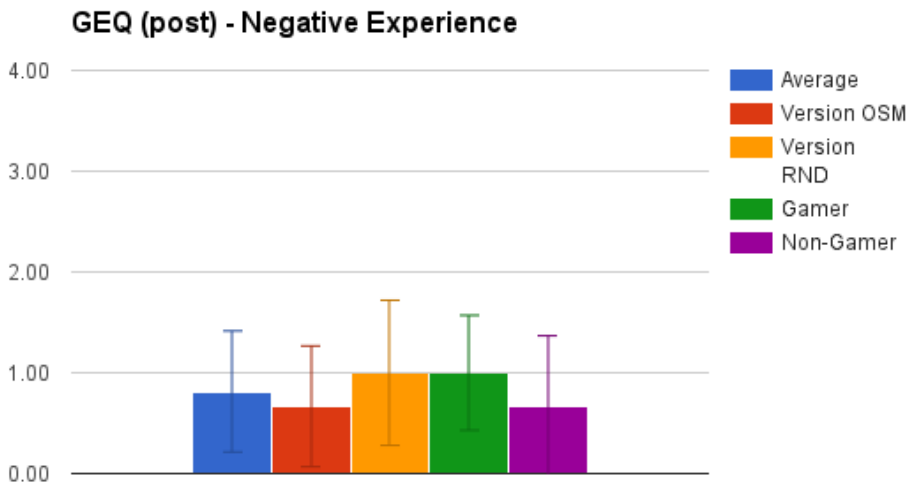

GEQ (core) - Flow
4.00
3.00
2.00
1.00
0.00
Average
Version OSM
Version RND
Gamer
Non-Gamer

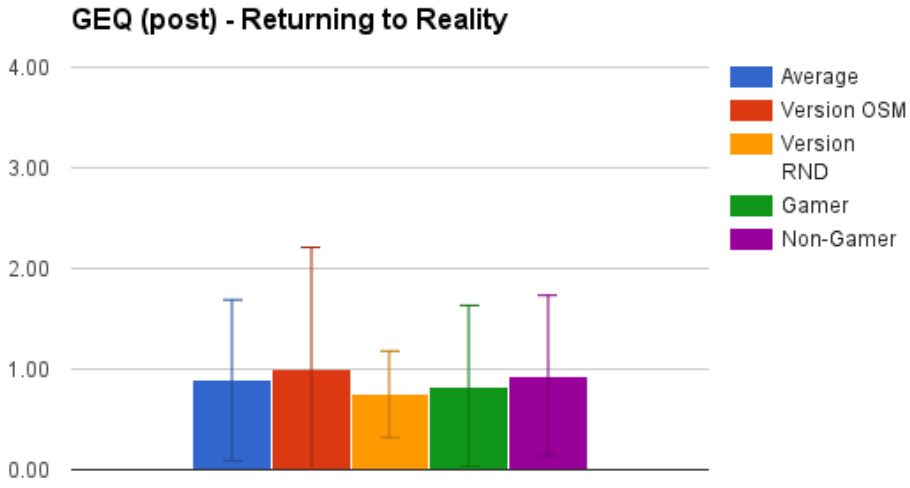

GEQ (core) - Challenge
4.00
3.00
2.00
1.00
0.00
Average
Version OSM
Version RND
Gamer
Non-Gamer

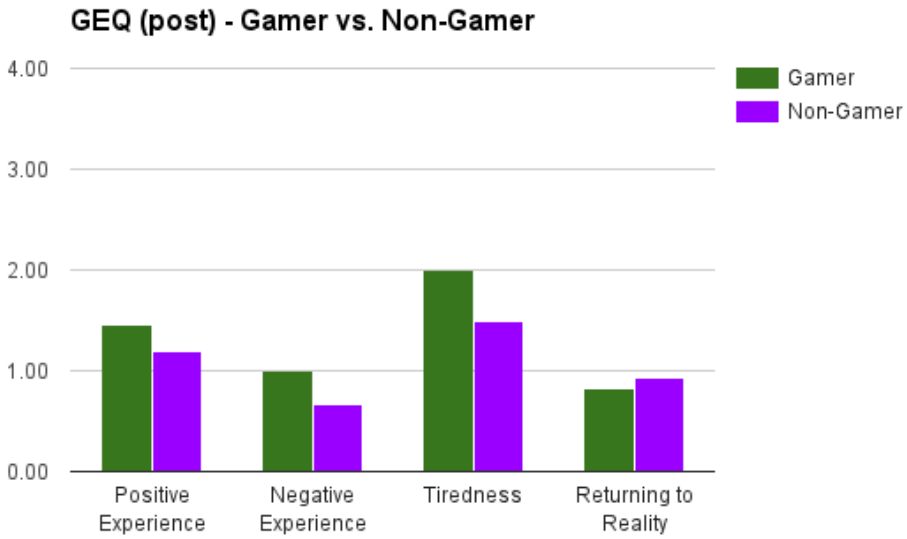

GEQ (core) - Negative affect
4.00
3.00
2.00
1.00
0.00
Average
Version OSM
Version RND
Gamer
Non-Gamer

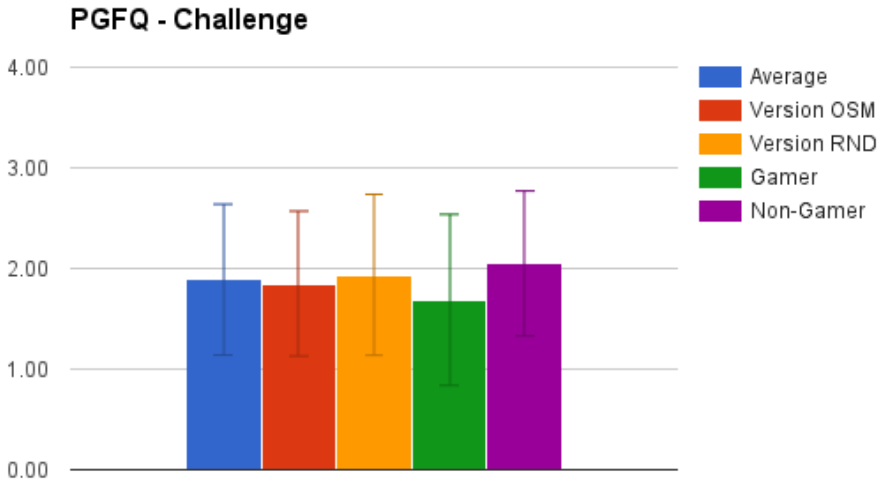

GEQ (core) - Positive affect
4.00
3.00
2.00
1.00
0.00
Average
Version OSM
Version RND
Gamer
Non-Gamer

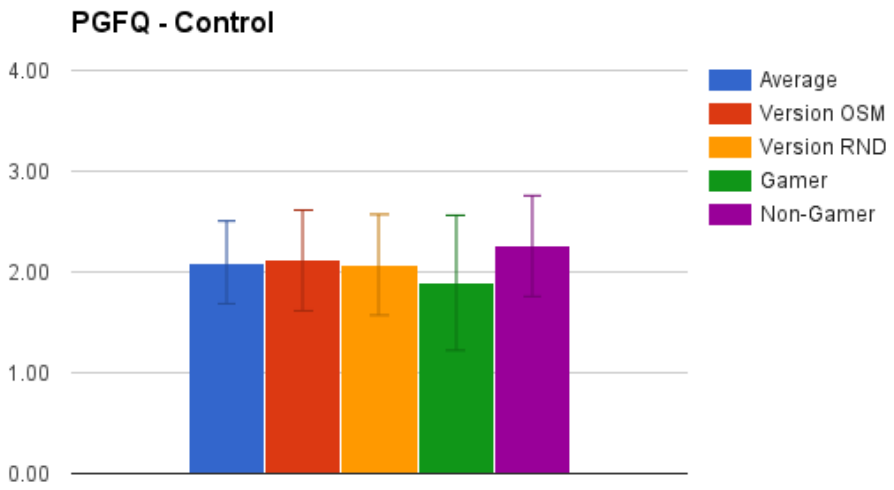

GEQ (core) - Gamer vs. Non-Gamer
4.00
3.00
2.00
1.00
0.00
Gamer
Non-Gamer
Competence
S. & I. Immersion
Flow
Tension/Annoyance
Challenge
Negative affect
Positive affect

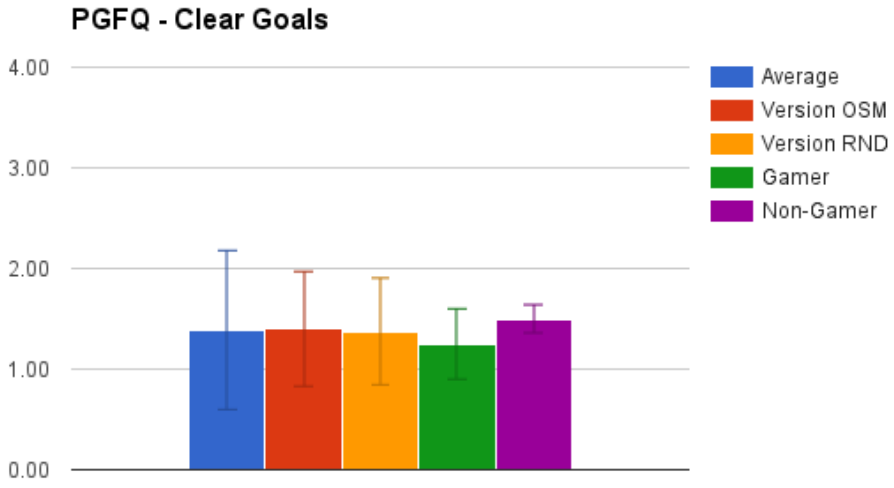

GEQ (post) - Negative Experience
4.00
3.00
2.00
1.00
0.00
Average
Version OSM
Version RND
Gamer
Non-Gamer

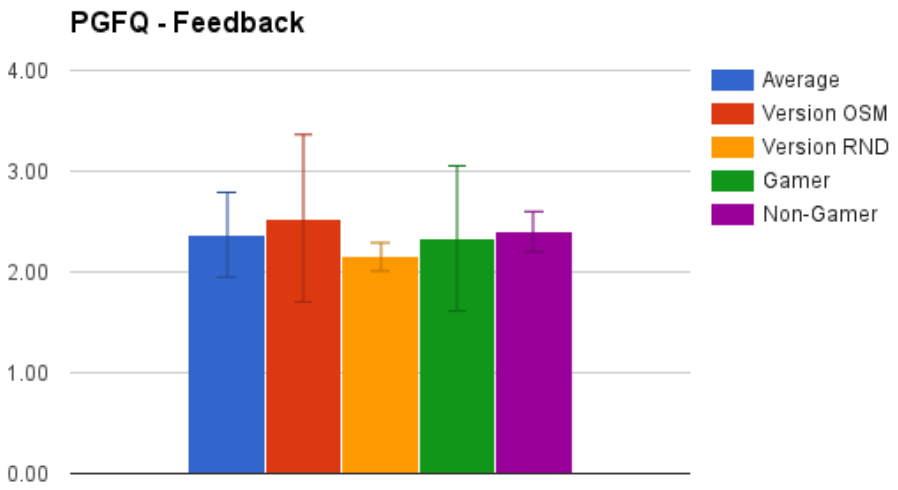

GEQ (post) - Returning to Reality
4.00
3.00
2.00
1.00
0.00
Average
Version OSM
Version RND
Gamer
Non-Gamer

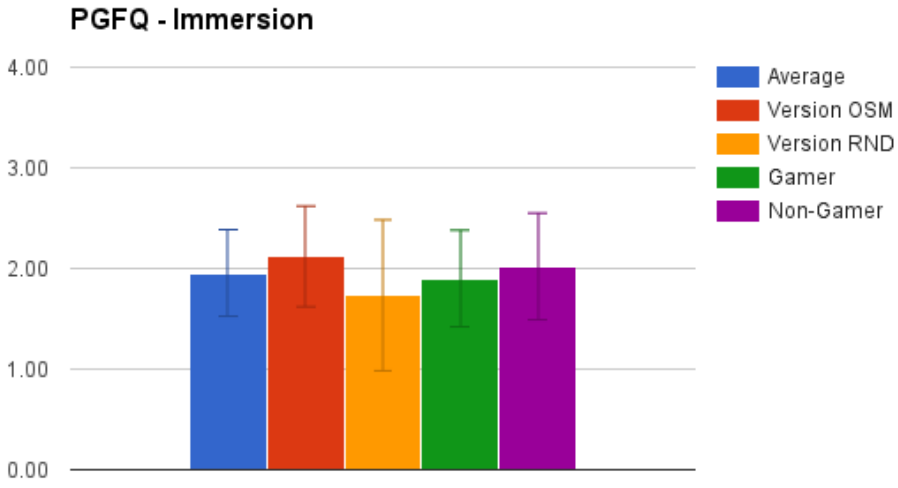

GEQ (post) - Gamer vs. Non-Gamer
4.00
3.00
2.00
1.00
0.00
Gamer
Non-Gamer
Positive Experience
Negative Experience
Tiredness
Returning to Reality

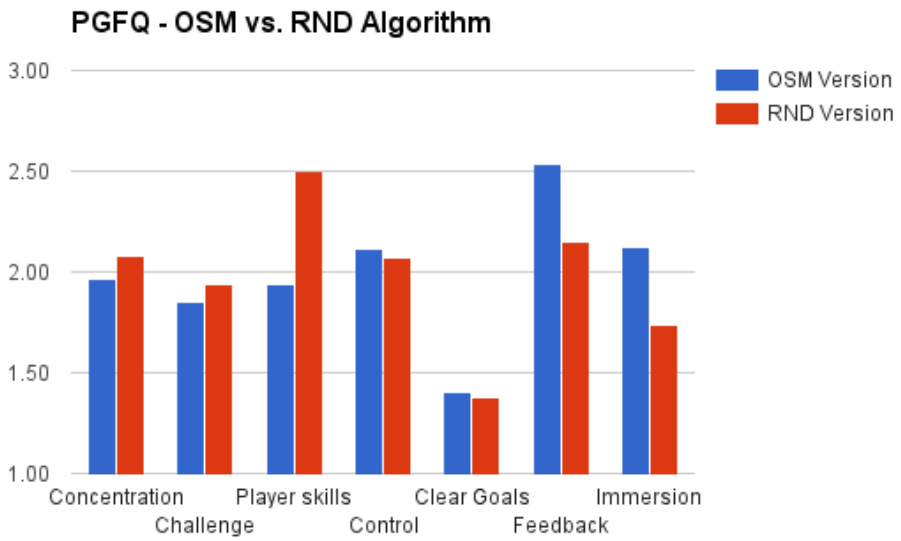

PGFQ - Challenge
4.00
3.00
2.00
1.00
0.00
Average
Version OSM
Version RND
Gamer
Non-Gamer

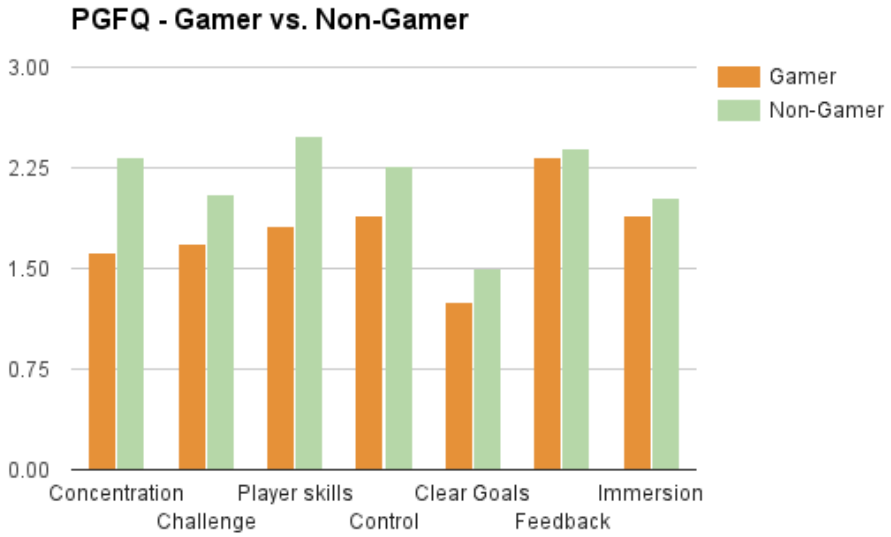

PGFQ - Control
4.00
3.00
2.00
1.00
0.00
Average
Version OSM
Version RND
Gamer
Non-Gamer

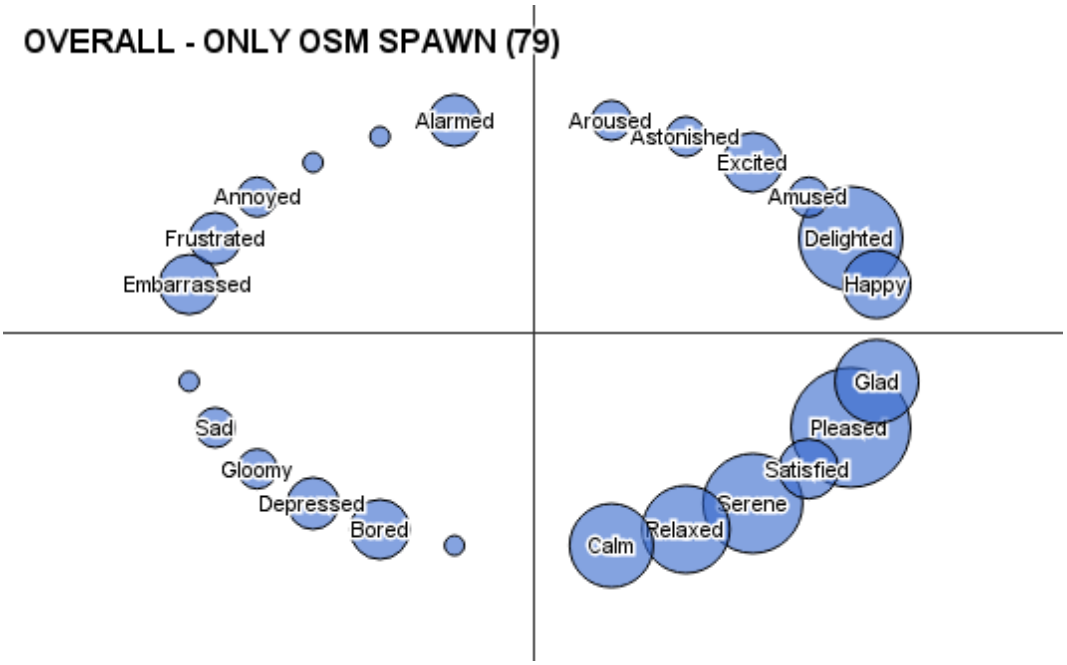

PGFQ - Clear Goals
4.00
3.00
2.00
1.00
0.00
Average
Version OSM
Version RND
Gamer
Non-Gamer

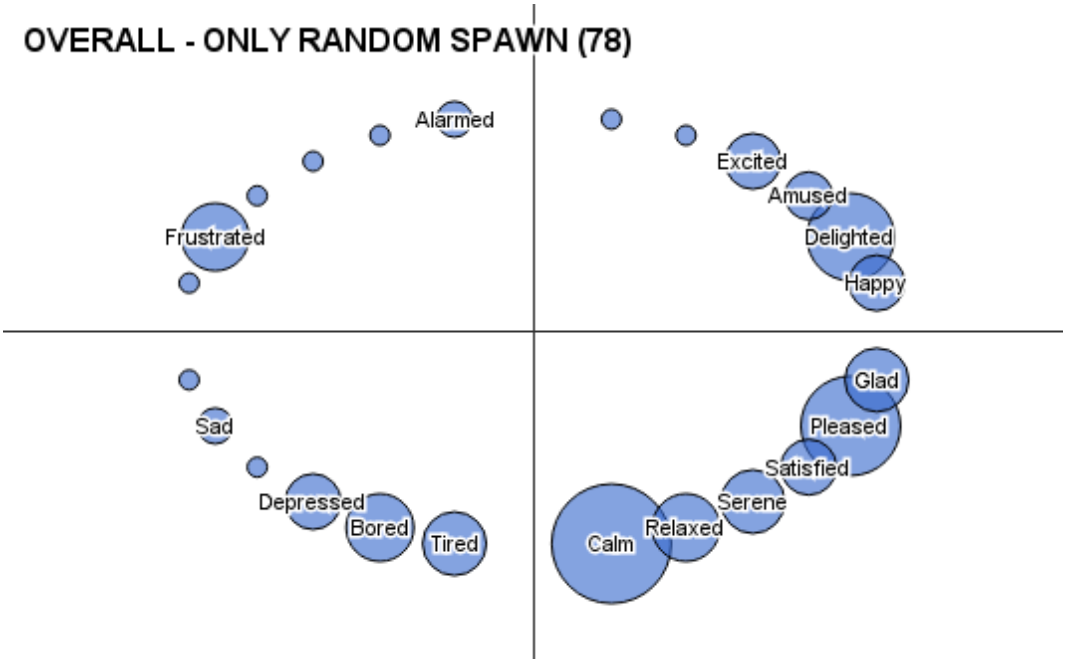

PGFQ - Feedback
4.00
3.00
2.00
1.00
0.00
Average
Version OSM
Version RND
Gamer
Non-Gamer

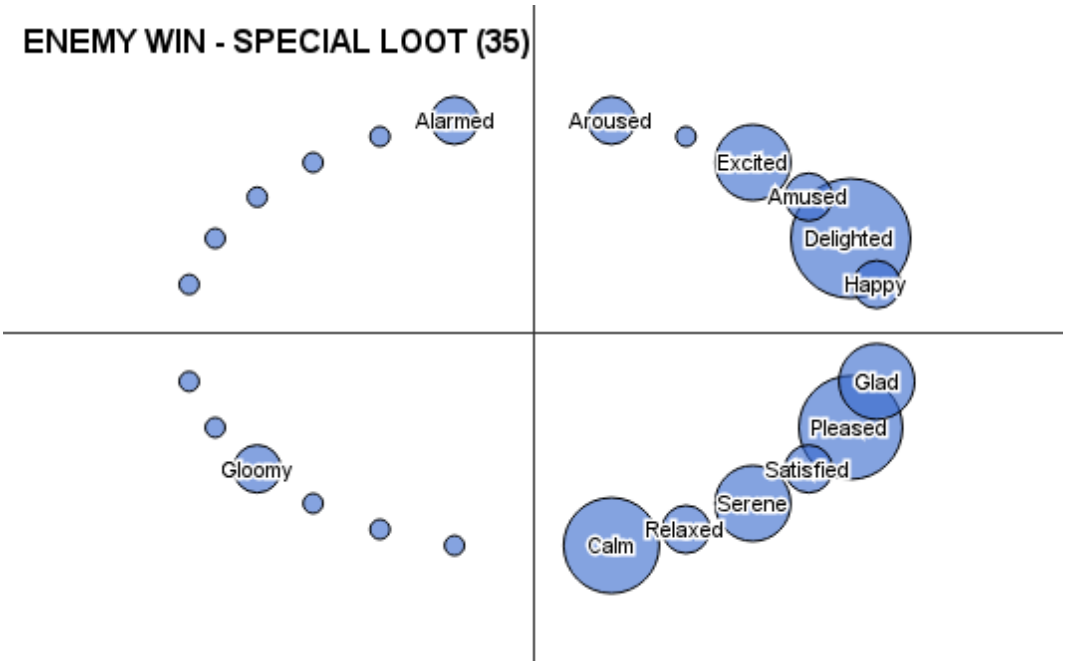

PGFQ - Immersion
4.00
3.00
2.00
1.00
0.00
Average
Version OSM
Version RND
Gamer
Non-Gamer

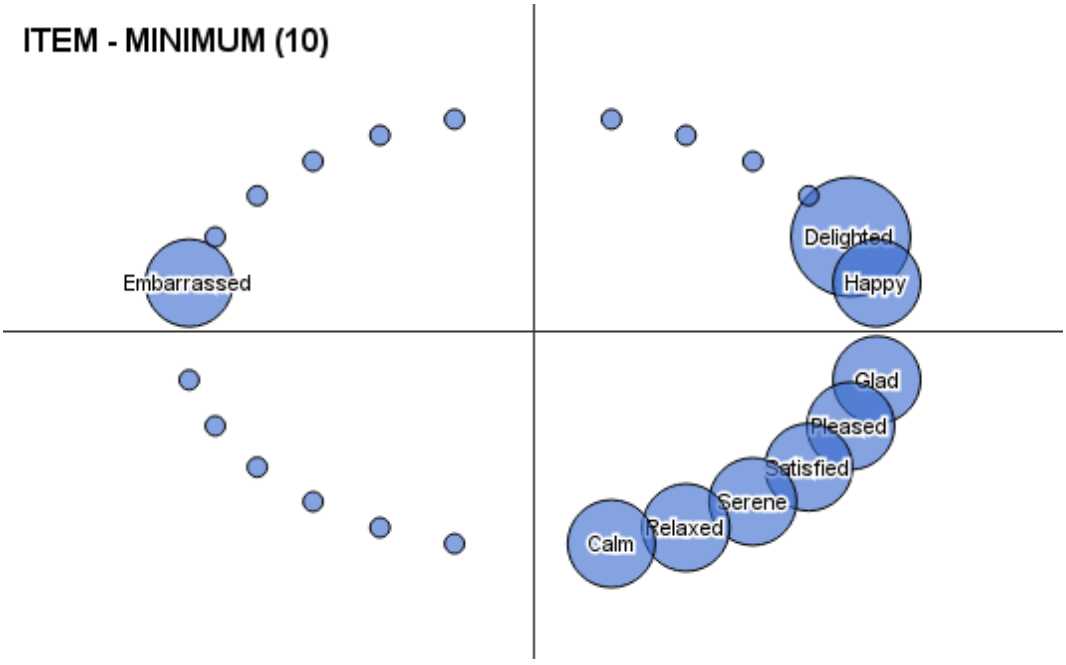

PGFQ - OSM vs. RND Algorithm
3.00
2.50
2.00
1.50
1.00
OSM Version
RND Version
Concentration
Challenge
Player skills
Control
Clear Goals
Feedback
Immersion

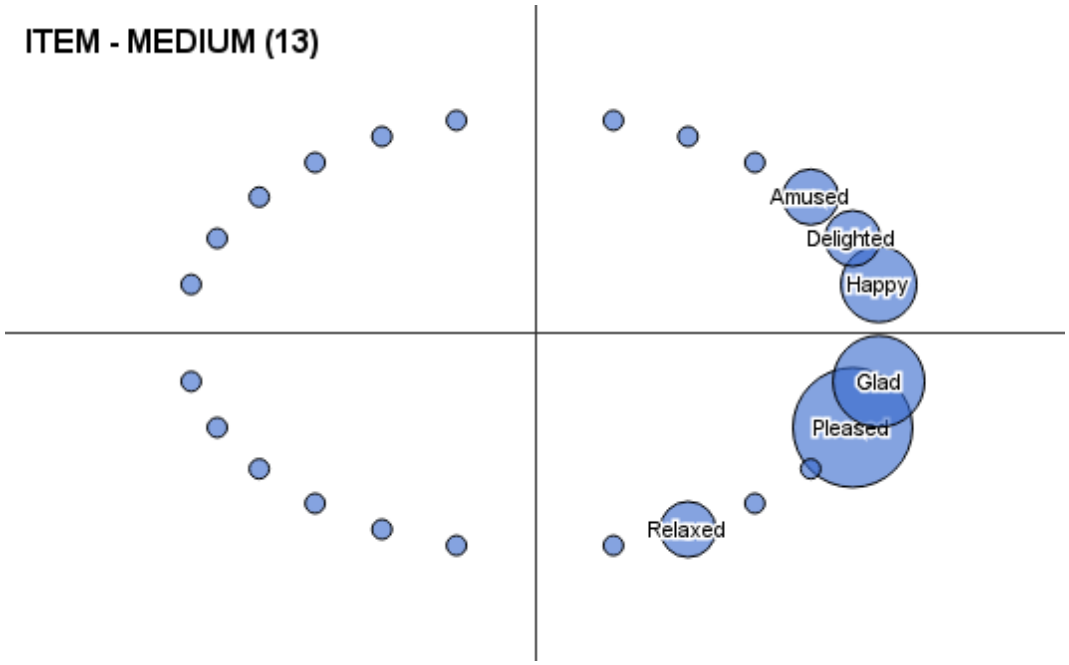

PGFQ - Gamer vs. Non-Gamer
3.00
2.25
1.50
0.75
0.00
Gamer
Non-Gamer
Concentration
Challenge
Player skills
Control
Clear Goals
Feedback
Immersion

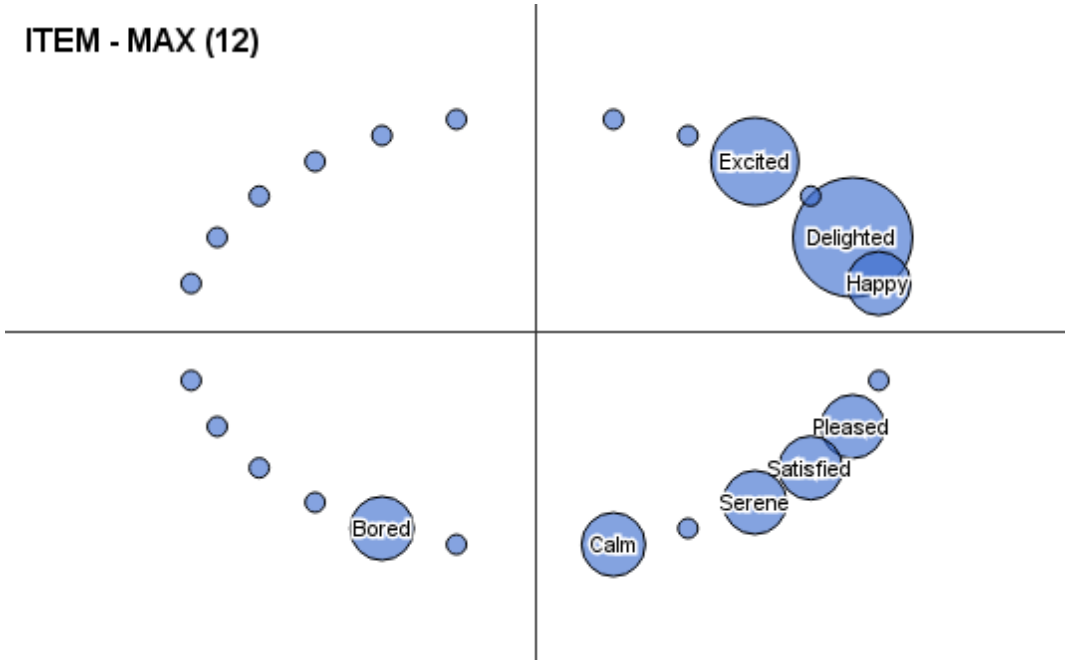

OVERALL - ONLY OSM SPAWN (79)
Alarmed
Annoyed
Frustrated
Embarrassed
Aroused
Astonished
Excited
Amused
Delighted
Happy
Sad
Gloomy
Depressed
Bored
Glad
Pleased
Satisfied
Serene
Relaxed
Calm


OVERALL - ONLY RANDOM SPAWN (78)
Alarmed
Frustrated
Excited
Amused
Delighted
Happy
Sad
Depressed
Bored
Tired
Glad
Pleased
Satisfied
Serene
Relaxed
Calm


ENEMY WIN - SPECIAL LOOT (35)
Alarmed
Aroused
Excited
Amused
Delighted
Happy
Gloomy
Glad
Pleased
Satisfied
Serene
Relaxed
Calm


ITEM - MINIMUM (10)
Embarrassed
Delighted
Happy
Glad
Pleased
Satisfied
Serene
Relaxed
Calm


ITEM - MEDIUM (13)
Amused
Delighted
Happy
Glad
Pleased
Relaxed


ITEM - MAX (12)
Excited
Delighted
Happy
Pleased
Satisfied
Serene
Calm
Bored

Bibliography

Charlene Jennett, Anna L. Cox, Paul Cairns, Samira Dhoparee, Andrew Epps, Tim Tijs, and Alison Walton. Measuring and defining the experience of immersion in games. *International Journal of Human-Computer Studies*, 66(9):641–661, September 2008. ISSN 1071-5819. doi: 10.1016/j.ijhcs.2008.04.004. URL http://www.sciencedirect.com/science/article/pii/S1071581908000499.

B. G. Witmer and M. J. Singer. Measuring Presence in Virtual Environments: A Presence Questionnaire. *Presence*, 7(3):225–240, June 1998. ISSN 1054-7460. doi: 10.1162/105474698565686.

Emily Brown and Paul Cairns. A Grounded Investigation of Game Immersion. In *CHI '04 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '04, pages 1297–1300, New York, NY, USA, 2004. ACM. ISBN 1-58113-703-6. doi: 10.1145/985921.986048. URL http://doi.acm.org/10.1145/985921.986048.

Timothy Sanders and Paul Cairns. Time Perception, Immersion and Music in Videogames. In *Proceedings of the 24th BCS Interaction Specialist Group Conference*, BCS '10, pages 160–167, Swinton, UK, UK, 2010. British Computer Society. ISBN 978-1-78017-130-2. URL http://dl.acm.org/citation.cfm?id=2146303.2146327.

Anna Cox, Paul Cairns, Pari Shah, and Michael Carroll. Not Doing but Thinking: The Role of Challenge in the Gaming Experience. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '12, pages 79–88, New York, NY, USA, 2012. ACM. ISBN 978-1-4503-1015-4. doi: 10.1145/2207676.2207689. URL http://doi.acm.org/10.1145/2207676.2207689.

Johanna Huhtala, Poika Isokoski, and Saila Ovaska. The Usefulness of an Immersion Questionnaire in Game Development. In *CHI '12 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '12, pages 1859–1864, New York, NY, USA, 2012. ACM. ISBN 978-1-4503-1016-1. doi: 10.1145/2212776.2223719. URL http://doi.acm.org/10.1145/2212776.2223719.

Penelope Sweetser and Peta Wyeth. GameFlow: A Model for Evaluating Player Enjoyment in Games. *Comput. Entertain.*, 3(3):3–3, July 2005. ISSN 1544-3574. doi: 10.1145/1077246.1077253. URL http://doi.acm.org/10.1145/1077246.1077253.

Mihaly Csikszentmihalyi. *Flow: The Psychology of Optimal Experience*. Harper & Row, New York, 1990. ISBN 978-0-06-016253-5.

Jeanne H. Brockmyer, Christine M. Fox, Kathleen A. Curtiss, Evan McBroom, Kimberly M. Burkhart, and Jacquelyn N. Pidruzny. The development of the Game Engagement Questionnaire: A measure of engagement in video game-playing. *Journal of Experimental Social Psychology*, 45(4):624–634, July 2009. ISSN 0022-1031. doi: 10.1016/j.jesp.2009.02.016. URL http://www.sciencedirect.com/science/article/pii/S0022103109000444.

Regina Bernhaupt, Manfred Eckschlager, and Manfred Tscheligi. Methods for Evaluating Games: How to Measure Usability and User Experience in Games? In *Proceedings of the International Conference on Advances in Computer Entertainment Technology*, ACE '07, pages 309–310, New York, NY, USA, 2007. ACM. ISBN 978-1-59593-640-0. doi: 10.1145/1255047.1255142. URL http://doi.acm.org/10.1145/1255047.1255142.

James Paul Gee. What Video Games Have to Teach Us About Learning and Literacy. *Comput. Entertain.*, 1(1):20–20, October 2003. ISSN 1544-3574. doi: 10.1145/950566.950595. URL http://doi.acm.org/10.1145/950566.950595.

Hua Qin, Pei-Luen Patrick Rau, and Gavriel Salvendy. Player Immersion in the Computer Game Narrative. In Lizhuang Ma, Matthias Rauterberg, and Ryohei Nakatsu, editors, *Entertainment Computing – ICEC 2007*, number 4740 in Lecture Notes in Computer Science, pages 458–461. Springer Berlin Heidelberg, 2007. ISBN 978-3-540-74872-4 978-3-540-74873-1. URL http:

//link.springer.com/chapter/10.1007/978-3-540-74873-1_60. DOI: 10.1007/978-3-540-74873-1_60.

Hua Qin, Pei-Luen Patrick Rau, and Gavriel Salvendy. Measuring Player Immersion in the Computer Game Narrative. *International Journal of Human-Computer Interaction*, 25(2):107–133, February 2009. ISSN 1044-7318. doi: 10.1080/10447310802546732. URL http://dx.doi.org/10.1080/10447310802546732.

Hua Qin, Pei-Luen Patrick Rau, and Gavriel Salvendy. Effects of different scenarios of game difficulty on player immersion. *Interacting with Computers*, 22(3):230–239, May 2010. ISSN 0953-5438. doi: 10.1016/j.intcom.2009.12.004. URL http://www.sciencedirect.com/science/article/pii/S095354380900109X.

Christiane Moser, Verena Fuchsberger, and Manfred Tscheligi. Rapid Assessment of Game Experiences in Public Settings. In *Proceedings of the 4th International Conference on Fun and Games*, FnG '12, pages 73–82, New York, NY, USA, 2012. ACM. ISBN 978-1-4503-1570-8. doi: 10.1145/2367616.2367625. URL http://doi.acm.org/10.1145/2367616.2367625.

Julian Frommel, Katja Rogers, Julia Brich, Daniel Besserer, Leonard Bradatsch, Isabel Ortinau, Ramona Schabenberger, Valentin Riemer, Claudia Schrader, and Michael Weber. Integrated Questionnaires: Maintaining Presence in Game Environments for Self-Reported Data Acquisition. In *Proceedings of the 2015 Annual Symposium on Computer-Human Interaction in Play*, CHI PLAY '15, pages 359–368, New York, NY, USA, 2015. ACM. ISBN 978-1-4503-3466-2. doi: 10.1145/2793107.2793130. URL http://doi.acm.org/10.1145/2793107.2793130.

Steve Benford, Carsten Magerkurth, and Peter Ljungstrand. Bridging the Physical and Digital in Pervasive Gaming. *Commun. ACM*, 48(3):54–57, March 2005a. ISSN 0001-0782. doi: 10.1145/1047671.1047704. URL http://doi.acm.org/10.1145/1047671.1047704.

Adrian David Cheok, Kok Hwee Goh, Wei Liu, Farzam Farbiz, Siew Wan Fong, Sze Lee Teo, Yu Li, and Xubo Yang. Human Pacman: A Mobile, Wide-area Entertainment System Based on Physical, Social, and Ubiquitous Computing. *Personal Ubiquitous Comput.*, 8(2):71–81, May

2004. ISSN 1617-4909. doi: 10.1007/s00779-004-0267-x. URL http://dx.doi.org/10.1007/s00779-004-0267-x.

Steve Benford, Duncan Rowland, Martin Flintham, Adam Drozd, Richard Hull, Josephine Reid, Jo Morrison, and Keri Facer. Life on the Edge: Supporting Collaboration in Location-based Experiences. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '05, pages 721–730, New York, NY, USA, 2005b. ACM. ISBN 1-58113-998-5. doi: 10.1145/1054972.1055072. URL http://doi.acm.org/10.1145/1054972.1055072.

Rod McCall, Shaleph O'Neil, and Fiona Carroll. Measuring Presence in Virtual Environments. In *CHI '04 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '04, pages 783–784, New York, NY, USA, 2004. ACM. ISBN 1-58113-703-6. doi: 10.1145/985921.985934. URL http://doi.acm.org/10.1145/985921.985934.

Marek Bell, Matthew Chalmers, Louise Barkhuus, Malcolm Hall, Scott Sherwood, Paul Tennent, Barry Brown, Duncan Rowland, Steve Benford, Mauricio Capra, and Alastair Hampshire. Interweaving Mobile Games with Everyday Life. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '06, pages 417–426, New York, NY, USA, 2006. ACM. ISBN 1-59593-372-7. doi: 10.1145/1124772.1124835. URL http://doi.acm.org/10.1145/1124772.1124835.

Steve Benford, Andy Crabtree, Stuart Reeves, Jennifer Sheridan, Alan Dix, Martin Flintham, and Adam Drozd. The Frame of the Game: Blurring the Boundary Between Fiction and Reality in Mobile Experiences. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '06, pages 427–436, New York, NY, USA, 2006a. ACM. ISBN 1-59593-372-7. doi: 10.1145/1124772.1124836. URL http://doi.acm.org/10.1145/1124772.1124836.

Steve Benford, Andy Crabtree, Martin Flintham, Adam Drozd, Rob Anastasi, Mark Paxton, Nick Tandavanitj, Matt Adams, and Ju Row-Farr. Can You See Me Now? *ACM Trans. Comput.-Hum. Interact.*, 13(1):100–133, March 2006b. ISSN 1073-0516. doi: 10.1145/1143518.1143522. URL http://doi.acm.org/10.1145/1143518.1143522.

Kenton O'Hara. Understanding Geocaching Practices and Motivations. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '08, pages 1177–1186, New York, NY, USA, 2008. ACM. ISBN 978-1-60558-011-1. doi: 10.1145/1357054.1357239. URL http://doi.acm.org/10.1145/1357054.1357239.

Josephine Reid. Design for Coincidence: Incorporating Real World Artifacts in Location Based Games. In *Proceedings of the 3rd International Conference on Digital Interactive Media in Entertainment and Arts*, DIMEA '08, pages 18–25, New York, NY, USA, 2008. ACM. ISBN 978-1-60558-248-1. doi: 10.1145/1413634.1413643. URL http://doi.acm.org/10.1145/1413634.1413643.

Steve Harrison and Paul Dourish. Re-place-ing Space: The Roles of Place and Space in Collaborative Systems. In *Proceedings of the 1996 ACM Conference on Computer Supported Cooperative Work*, CSCW '96, pages 67–76, New York, NY, USA, 1996. ACM. ISBN 0-89791-765-0. doi: 10.1145/240080.240193. URL http://doi.acm.org/10.1145/240080.240193.

Paul Dourish. Re-space-ing Place: "Place" and "Space" Ten Years on. In *Proceedings of the 2006 20th Anniversary Conference on Computer Supported Cooperative Work*, CSCW '06, pages 299–308, New York, NY, USA, 2006. ACM. ISBN 1-59593-249-6. doi: 10.1145/1180875.1180921. URL http://doi.acm.org/10.1145/1180875.1180921.

Iris Soute, Maurits Kaptein, and Panos Markopoulos. Evaluating Outdoor Play for Children: Virtual vs. Tangible Game Objects in Pervasive Games. In *Proceedings of the 8th International Conference on Interaction Design and Children*, IDC '09, pages 250–253, New York, NY, USA, 2009. ACM. ISBN 978-1-60558-395-2. doi: 10.1145/1551788.1551844. URL http://doi.acm.org/10.1145/1551788.1551844.

Annika Waern, Markus Montola, and Jaakko Stenros. The Three-sixty Illusion: Designing for Immersion in Pervasive Games. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '09, pages 1549–1558, New York, NY, USA, 2009. ACM. ISBN 978-1-60558-246-7. doi: 10.1145/1518701.1518939. URL http://doi.acm.org/10.1145/1518701.1518939.

Tara Carrigy, Katsiaryna Naliuka, Natasa Paterson, and Mads Haahr. Design and Evaluation of Player Experience of a Location-based Mobile Game. In *Proceedings of the 6th Nordic Conference on Human-Computer Interaction: Extending Boundaries*, NordiCHI '10, pages 92–101, New York, NY, USA, 2010. ACM. ISBN 978-1-60558-934-3. doi: 10.1145/1868914.1868929. URL http://doi.acm.org/10.1145/1868914.1868929.

Kate Lund, Mark Lochrie, and Paul Coulton. Enabling Emergent Behaviour in Location Based Games. In *Proceedings of the 14th International Academic MindTrek Conference: Envisioning Future Media Environments*, MindTrek '10, pages 78–85, New York, NY, USA, 2010. ACM. ISBN 978-1-4503-0011-7. doi: 10.1145/1930488.1930505. URL http://doi.acm.org/10.1145/1930488.1930505.

Juha Tiensyrjä, Timo Ojala, Toni Hakanen, and Ossi Salmi. panOULU Conqueror: Pervasive Location-aware Multiplayer Game for City-wide Wireless Network. In *Proceedings of the 3rd International Conference on Fun and Games*, Fun and Games '10, pages 157–165, New York, NY, USA, 2010. ACM. ISBN 978-1-60558-907-7. doi: 10.1145/1823818.1823836. URL http://doi.acm.org/10.1145/1823818.1823836.

Jason Kurczak, T. C. Nicholas Graham, Claire Joly, and Regan L. Mandryk. Hearing is Believing: Evaluating Ambient Audio for Location-based Games. In *Proceedings of the 8th International Conference on Advances in Computer Entertainment Technology*, ACE '11, pages 32:1–32:8, New York, NY, USA, 2011. ACM. ISBN 978-1-4503-0827-4. doi: 10.1145/2071423.2071463. URL http://doi.acm.org/10.1145/2071423.2071463.

María Guadalupe Alvarez Díaz, Marcus Toftedahl, and Torbjörn Svensson. The Mystery of Elin. Incorporating a City Cultural Program on History and Heritage into a Pervasive Game. In *Proceedings of the 2014 Conference on Interactive Entertainment*, IE2014, pages 9:1–9:10, New York, NY, USA, 2014. ACM. ISBN 978-1-4503-2790-9. doi: 10.1145/2677758.2677768. URL http://doi.acm.org/10.1145/2677758.2677768.

Jason Procyk and Carman Neustaedter. GEMS: The Design and Evaluation of a Location-based Storytelling Game. In *Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing*,

CSCW '14, pages 1156–1166, New York, NY, USA, 2014. ACM. ISBN 978-1-4503-2540-0. doi: 10.1145/2531602.2531701. URL http://doi.acm.org/10.1145/2531602.2531701.

Paul Spiesberger, Florian Jungwirth, Christoph Wöss, Stefan Bachl, Johannes Harms, and Thomas Grechenig. Woody: A Location-based Smartphone Game to Increase Children's Outdoor Activities in Urban Environments. In *Proceedings of the 14th International Conference on Mobile and Ubiquitous Multimedia*, MUM '15, pages 368–372, New York, NY, USA, 2015. ACM. ISBN 978-1-4503-3605-5. doi: 10.1145/2836041.2841210. URL http://doi.acm.org/10.1145/2836041.2841210.

Anne E. Bowser, Derek L. Hansen, Jocelyn Raphael, Matthew Reid, Ryan J. Gamett, Yurong R. He, Dana Rotman, and Jenny J. Preece. Prototyping in PLACE: A Scalable Approach to Developing Location-based Apps and Games. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '13, pages 1519–1528, New York, NY, USA, 2013. ACM. ISBN 978-1-4503-1899-0. doi: 10.1145/2470654.2466202. URL http://doi.acm.org/10.1145/2470654.2466202.

Carman Neustaedter, Anthony Tang, and Tejinder K. Judge. Creating Scalable Location-based Games: Lessons from Geocaching. *Personal Ubiquitous Comput.*, 17(2):335–349, February 2013. ISSN 1617-4909. doi: 10.1007/s00779-011-0497-7. URL http://dx.doi.org/10.1007/s00779-011-0497-7.

Wijnand IJsselsteijn, Yvonne de Kort, Karolien Poels, Audrius Jurgelionis, and Francesco Bellotti. Characterising and Measuring User Experiences in Digital Games. *Proceedings of the International Conference on Advances in Computer Entertainment Technology (ACE 2007)*, June 2007.

Kalle Jegers. Pervasive Game Flow: Understanding Player Enjoyment in Pervasive Gaming. *Comput. Entertain.*, 5(1), January 2007. ISSN 1544-3574. doi: 10.1145/1236224.1236238. URL http://doi.acm.org/10.1145/1236224.1236238.

Jesse Schell. *The Art of Game Design: A Book of Lenses*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2008. ISBN 0-12-369496-5.

Lynne Baillie, Lee Morton, David C. Moffat, and Stephen Uzor. Capturing the response of players to a location-based game. *Personal Ubiquitous Comput.*, 15(1):13–24, January 2011. ISSN 1617-4909. doi: 10.1007/s00779-010-0309-5. URL http://dx.doi.org/10.1007/s00779-010-0309-5.

Kiel M Gilleade and Alan Dix. Using Frustration in the Design of Adaptive Videogames. In *Proceedings of the 2004 ACM SIGCHI International Conference on Advances in Computer Entertainment Technology*, ACE '04, pages 228–232, New York, NY, USA, 2004. ACM. ISBN 1-58113-882-2. doi: 10.1145/1067343.1067372. URL http://doi.acm.org/10.1145/1067343.1067372.