Applying Usability Testing Methods into Game Development: Case Casters of Kalderon

University of Oulu
Department of Information Processing Science
Master's Thesis
Panu Pasanen
16.4.2016

Abstract

Small game development companies don't nessecarily have strict prodecures or methods when it comes to testing their product. This was the case with the Oulu based game company Meizi Games, who felt the need to conduct a testing activity for their still developed game, Casters of Kalderon.

Casters of Kalderon was tested through conducting usability testing in order to find out the issues it currently has. These issues would be communicated to the developers by conducting the testing with potential customers eg. People who play mobile games. The testing activities were observed and feedback was gathered, and the results were delivered to and discussed with the developers of the game.

As Casters of Kalderon is a relatively complex strategy game, the biggest usability issue it has is the lacking in introducing the game's many mechanics to new players. The game currently has a tutorial mechanism that aims to help the player get familiar with the game, but it is currently poorly implemented, as the test results show. This would often cause confusion and frustration in the testers, who had no previous experience with the game. There were also some usability related issues that came in the way of the players being able to enjoy the game.

Game developers need to pay close attention to the designing of tutorials for complex games, like this case. Meizi Games found the usability testing activities to be a feasible way to find out the flaws and issues in their game, and promised to work on the found issues to improve the game in the future.

Keywords
Playability, Game Flow, Player Engagement, Usability Testing

Supervisor
PhD, university lecturer Mikko Rajanen

Foreword

I would like to thank PhD Mikko Rajanen for the supervision and open mindedness when it came to the research subject, as it was quite a difficult topic to grasp. I would also like to thank all the testing participants, as well as the people at Meizi Games for making this thesis possible.

Panu Pasanen

Oulu, April 16, 2016

Contents

Abstract 1
Foreword 2
Contents 3
1. INTRODUCTION 4
2. BACKGROUND 5
    2.1 Playability 5
    2.2 Game Flow 6
    2.3 Usability Testing 6
3. RESEARCH MATERIALS & METHODS 8
    3.1 Case: Casters of Kalderon 8
    3.2 Previous Testing Results 9
    3.3 Research Method 12
4. RESULTS 13
    4.1 Testing session 1 in Lab 13
    4.2 Testing session 2 in Lab 17
    4.3 Remote testing 20
    4.4 Issue prioritization 21
    4.5 Limitations & Benefits 22
5. CONCLUDING DISCUSSION 23
References 26

1. INTRODUCTION

The concept of usability has many names when it comes to research related to digital games. “Playability”, “gameplay” & “player experience” are all concepts that are somehow connected to the usability of a game presented in scientific literature related to games. According to the ISO 9241-11 standard, usability is defined as “the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use” (Marghescu, 2009). The ISO/IEC 9126-1 standard defines usability as “the capability of the software product to be understood, learned, used and attractive to the user, when used under specified conditions” and quality in use as “the capability of the software product to enable specified users to achieve specified goals with effectiveness, productivity, safety and satisfaction in specified contexts of use (Marghescu, 2009). ISO 9241-11 has a focus on describing the context of use, but does not go into detail with the different phases, activities and inputs and outputs required in the evaluation process. ISO/IEC 9126-1 provides more extensive models of usability in use. (Marghescu, 2009).

However, experiencing digital games is not the same thing as experiencing the more traditional software products. Playability, according to Nacke (2009), is defined as the game’s *“capability to provide enjoyment for a player over an extended period of time”.* On the other hand, Järvinen & Mäyrä (2002) define playability as *“a collection of criteria with which to evaluate a product’s gameplay or interaction”,* while also introducing four components of playability: functional, structural, audiovisual and social playability. The game usability and the quality of user interface in games play an important role in game playing experience and it is an important factor in the decision of buying a game (Rajanen & Marghescu 2006).

In this thesis the concept of player engagement is approached through conducting a usability testing workshop with users who are already familiar with gaming. The aim is to find out the key factors that keep players coming back to the same game over and over again (eg. the term engagement). (Schoenau-Fog, 2011) The workshop for the testing is done in cooperation with a Finnish game company with a game currently under development. The developers wish to conduct this usability testing study in order to find out the good and bad in their game from a potential customer point-of-view and identify possible issues with the current game, in order to make improvements on it. The results of this research provide interesting points for future game development about the issues that need to be taken into consideration when designing a game of similar nature. The results also show how similar testing techniques could be used to conduct game testing, even at the very late phases of the development.

The research problem for this thesis also stems from this concept of player engagement. The research problem of this thesis is “How do usability issues affect the enjoyment of players in this case.” With supporting questions “What are the issues identified and Can they be generalized into other games of the same genre?”

2. BACKGROUND

This chapter introduces the concepts of playability, game flow and usability, as these concepts provide a good background for the issues examined in this thesis.

2.1 Playability

Games are meant to be an interactive experience that focuses on player progression inside the game system, rather than the outcome obtained by playing the game. (Nacke, Schild & Niesenhaus, 2010). Games are played for the experience, which means that this experience needs to be well designed. In traditional software, the approach is often to design a pleasant way to a achieve a set goal in the most efficient way. This means that the development practices in the game industry must take this design aspect into account. Concepts like *flow* (Chen, 2007, Cowley et al, 2008), *immersion* (Sweetser & Wyeth, 2005) and *enjoyment* (Sweetser & Wyeth, 2005) are the building blocks for the game experience, which can be facilitated by choices in the game's design, to make it an *enjoyable, interactive* and *goal-driven experience,* and the underlying technical prerequisites, like the *user interface* and the *device platform* (Nacke, Schild & Niesenhaus, 2010). The experience the player gains from the game comes from the relationship between the player and the game. An important difference between playability and usability in that entertainment outweighs productivity as the primary motivating factor for the game-play activity (Cowley et al, 2008).

Game developers aim for their game to achieve long term engagement in the players. It is defined by Febretti & Garzotto as "the degree of voluntary use of a system along a wide period of time (weeks, months, even years), involving dozens, if not thousands, of interactions, each one spanning for significantly longer than few seconds or minutes." (Febretti & Garzotto, 2009 & Schoenau-Fog, 2011). Typical game sessions in a long game usually last one hour or more. However games are often designed to last for tens of hours, some even hundreds, or even thousands of hours for the overall experience (eg. Massive Multiplayer Online Games). A large amount of users repeatedly using a game for significantly longer than few minutes, while still returning after weeks, months or even years of playing it, indicates user satisfaction and make the game more reliable, which contributes to the quality of the game experience. (Febretti & Garzotto, 2009).

As games present an interface between the user and the game itself, usability can be considered as a key factor, when it comes to user engagement. Usability problems are expected to reduce the long term engagement of players, according to Febretti & Garzotto (2009). If usability issues come in the way of player enjoyment (the *fun* of the game), they are very likely to not return to it. This means that the developers need to make sure that the players are able to reach the fun of the game with the least amount of obstacles.

In their study Febretti & Garzotto (2009) present that engagement appears to be related significantly to the playability factors in game experiences, but there is only a low correlation between usability and long term engagement in the same context. The decrease in engagement due to usability problems was found to be only momentary. Unless the usability issue was really hindering, the users found new ways to overcome the usability problems and still continued playing (Febretti & Garzotto, 2009). These results show that users can be quite adaptable when playing games and at the end of the

play session, users declared themselves to be satisfied anyway, and would return to the game, despite the minor issues with usability. This is also backed up by previous research by Norman (2002) and Brown & Cairns (2004).

Febretti & Garzotto (2009) also conclude that games with shorter play sessions (eg. casual games) would manifest usability defects more frequently, due to the simpler gameplay structure, so weaknesses in usability could have a much larger impact on the overall engagement of the players and that similar outcomes could come from games developed by small non-professional teams, or games that do not go through a consistent Quality Assessment process. (Febretti & Garzotto, 2009).

2.2 Game Flow

In the context of games, immersion is very important to the player's enjoyment of the game (Brown & Cairns, 2004). The concept of immersion is broad, and while it seems to be understood by the gaming community, the definition and the cause of immersion is not exactly clear.

Previous study conducted by Brown & Cairns (2004) propose three levels of player immersion: engagement, engrossment and total immersion. Immersion is not seen as a static experience but is seen as a scale of involvement with the game. It was found not to be a necessary feature for enjoyment, but on the other hand immersion is always seen as an enjoyable experience. It is also noted in previous research that immersion and engagement take a hit from usability and control problems in games (Brown & Cairns, 2004). They conclude that there needs to be an invisibility of controls in order to achieve total immersion, which issues in usability can hinder. The player needs to feel like he or she is actually present in the game world with nothing obstructing their interactions with the game. (Brown & Cairns, 2004).

Flow in games has a lot to do with immersion, as it is described in previous literature as something that alters the user's sense of time and sense of self (Chen, 2007). In order for the player to be totally immersed in the game, they need to achieve this so called "flow state". (Brown & Cairns, 2004). As immersion has a lot to do with how much the user is enjoying the game, game developers need to find ways to actually see if their game is what the potential customers are looking for. It is crucial to point out any issues that may hinder the player experience, as there is always the potential of something in the game breaking the flow and the immersion of the player, in doing so removing the fun from the game, and possibly driving off potential customers. In order to find out what players think about the game, and how they play it, developers most often turn to activities that measure the usability of their game. (Rajanen & Nissinen, 2015).

2.3 Usability Testing

To better understand player experience, game developers need to get into close contact with the potential "customers", the players. The more traditional software projects usually have four main objectives: produce the required functionality, within budget and schedule while meeting an acceptable level of quality (Kasurinen & Smolander, 2014). However, can these objectives be applied into game development? Kasurinen & Smolander (2014) found in their study, that game developers tend to focus on the more "softer" values of development, like the actual content and the user experience, instead of reliability or efficiency. In software testing the objective is often to verify and validate the product under development, to see that it is done according to design and that it fulfills its purpose. Game development is based on software development, but some requirements are added, like the visual representation (the graphics) and creative designs with the game mechanics. How do game developers test these aspects then?

Usability testing is a testing activity that is conducted to improve the usability of a software product. Another main goal is to also improve the process of the design and development, so that issues that were previously apparent, can be avoided in the future with other products. This distinguishes usability testing from a research study, where the goals is to examine the existence of a phenomenon. (Dumas & Redish, 2009) Usually in usability testing, the tasks that can be done using the product are tested using the potential end users of the product. The participants of the test are asked for their opinions on the product by filling a questionnaire or by conducting an interview. (Dumas & Redish, 2009) In usability testing, the person conducting the test observes the participants while making notes and recordings on both their performance with the product and their comments. Usability testing is successful only if it helps to improve the tested product and its development process. (Dumas & Redish, 2009)

Kasurinen & Smolander (2014) found in their study, that most game developers consider their work to be creative, and do not consider themselves to strictly work in the business of software. Even the developers doing programming and testing considered these activities more creative, than in traditional software development. Another interesting finding they present, is that in game development, the process of testing has a much larger influence on the actual product, than in the more traditional software development. This comes from the fact that late changes are allowed and even expected to happen when developing games. Test results can result even in large revisions of the game, and major features could be changed based on the results of testing. This is mostly due to the developer's decision to test every aspect of the game with a differentiated test audience. In all of the game companies that took part in the research, the features that the intended audience did not like, were removed, revised or redesigned (Kasurinen & Smolander, 2014).

According to Kasurinen & Smolander (2014), the largest and most influential testing activity in game development appears to be usability testing, with an assigned test group consisting of the target demographic and other volunteers. The applied test approaches rely heavily on the usability, not on strictly documented test plans or test cases. The major focus of game companies, when it comes to testing, seems to be the "softer" aspects, like the user experience, usability and game mechanics, and not so much the technical aspects, unlike in traditional software development. In game development the user experience is seen to be in a more important role, which would support the earlier definition given by Nacke et. al, that games are meant to be "experiences" for the player. The popularity of usability testing in games is in contrast with game usability heuristics, which are popular among game usability researchers (e.g. Desurvire, H., & Wiberg, C. (2009), Desurvire, H., Caplan, M., & Toth, J. A. (2004), Pinelle, D., Wong, N., & Stach, T. (2008), Federoff, M. A. (2002), Pinelle, D., Wong, N., Stach, T., & Gutwin, C. (2009)) but according to a recent game usability survey in Northern European game companies, heuristic evaluation is not actually used by these game companies (Rajanen & Nissinen 2015).

3. RESEARCH MATERIALS & METHODS

In this chapter the research material is presented. As the research was done on an already existing game product, the concepts and problem areas of that game are introduced. The research method is also discussed in this chapter.

3.1 Case: Casters of Kalderon

Casters of Kalderon is a large-scale strategy game with an emphasis on managing a village, it's resources and battling against enemy armies, both player controlled and non-playable characters (NPCs). The game also has a focus on social aspects in form of so called "coalitions" that provide a way for the players to form unions in order to both give and receive help in times of need. The player starts with an empty village and then begins to expand it with buildings that serve different functions, such as farms for food, mines for minerals and a barracks for recruiting soldiers for battle. The game is under development for both mobile and pc platforms, with an emphasis on the latter, with a free-to-play monetization model that enables the player to speed up the different tasks by paying a small amount of money, a so called "micro transaction". These micro transactions are handled with a resource called "moonstones" that the player can use to place buildings into a construction queue. Once placed the buildings can then be upgraded, which can be sped up by using these moonstones.

Figure 1. Casters of Kalderon, A view of the player's village, with the GUI

The game introduces its many functions to the player by using tutorials. Tutorials in games aim to provide the player the very basics of the game, from which the player can explore and expand their mastery upon. As pointed out by Andersen et al. (2012), teaching the mechanics of the game can be a challenging task for the developers, but it is a very crucial process for engaging the players and keeping them interested in playing the game. Game developers often utilize tutorials in order to aid the player's learning, and are often the very first part of the game. This poses an emphasis on effective design of the tutorials, which is often based only on intuition (Andersen et. al, 2012).

A more deeper understanding of how tutorials impact the engagement of players would really help the developers to create better interfaces. Andersen et. al (2012) conclude that the effectiveness of using tutorials depends on the complexity of the game. They also conclude that it is unlikely that a single approach to tutorials will be a good fit for all kinds of games. Since players are explorative by nature, it is important to design the early phases of the game in a way that helps maximize the player's freedom to experiment and discovers the game's mechanics. Andersen et. al (2012) found little evidence that restricting players freedom to focus on a particular part of the interface or mechanic of the game is actually beneficial. As Casters of Kalderon is a large-scale game, with multiple important functions in order to enjoy it to the fullest extent, these tutorials play a very important part in introducing the player to the game. If they are not handled correctly, they may cause confusion in the players, which leads to frustration, which in turn makes them abandon the game.

3.2 Previous Testing Results

The game is a long way into development, with plenty of elements and assets already in place and it closely resembles a finished product. Most of the graphical objects, functionality and the GUI is already done. Some of the planned features, like the campaign mode that presents the player with different scenarios to complete, are still under development. The game has previously been tested with different game market analytics and metrics used from january 2014 to march 2014, and what they show tells a story of player engagement not forming. The amount of players that would return to the game after first starting it up, was 59,4% and the amount of users went down from 16 112 to 9 563 after the first start.

The game's introduction consists of going through tutorial paths, each consisting of a different amount of steps the player must complete in order to pass the tutorial. First the purpose of the tutorial is explained, then the steps within it need to be completed by the player and finally the player gets to claim their reward for completing the tutorial.

The developers provided statistics on how the users got through these tutorials provided in the game (see **Figure 2** and **Figure 3**). Out of the 10 294 users measured, only 4 446 went with the tutorials. This would suggest that less than half of the players had any interest in doing these tutorials. In turn this would suggest that these tutorials are either difficult to complete, or there are obstructions within the game that make players feel like they do not understand the point of these tutorials. The testing process aims to find out and pinpoint these moments, where the players possibly feel confused about the tutorials.

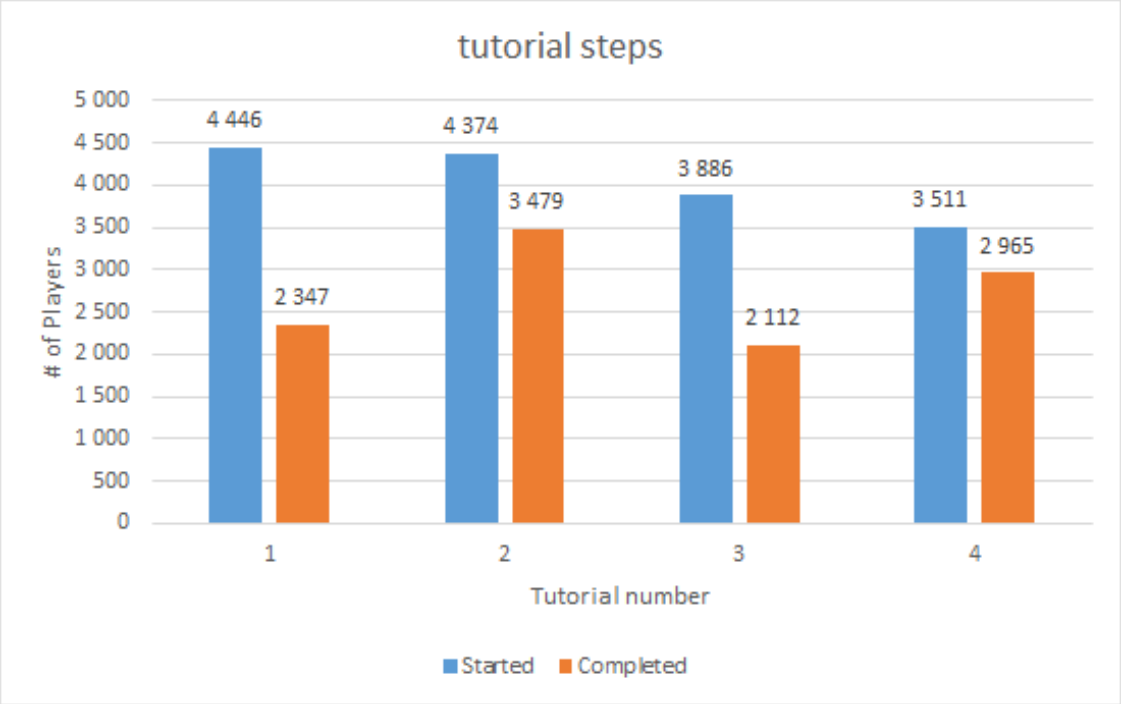

Figure 2. Tutorial steps started & completed.

The number of completed users represent those users that claimed the prize provided by the game after successfully finishing the set of tasks provided by the tutorial. As can be seen from the figure, roughly half of the players that started the first tutorial path, did not claim the prize provided by completing it. This would suggest that it was either hard to find or to understand, that the players are actually rewarded for completing these tutorials.

As players progress with the tutorial steps, the number of players that start the tutorial step decreases. This would suggest that as the player progresses in the game, their attention and interest to completing these tutorials starts to decline. Although in the last tutorial, 2 965 of the 3 511 players that started it, actually finished it. This could indicate that the more motivated players actually wanted to go through with these tutorials to learn the rules of the game, while those with a lesser span of attention, started to move away from the tutorials after the first step, and wanted to go on their way and experiment with the game. These results support that using the tutorials as they are now is not an effective way to engage new players to the game. Another factor highlighting this, is the retention numbers provided by the previous testing activities done with the game (Figure 2).

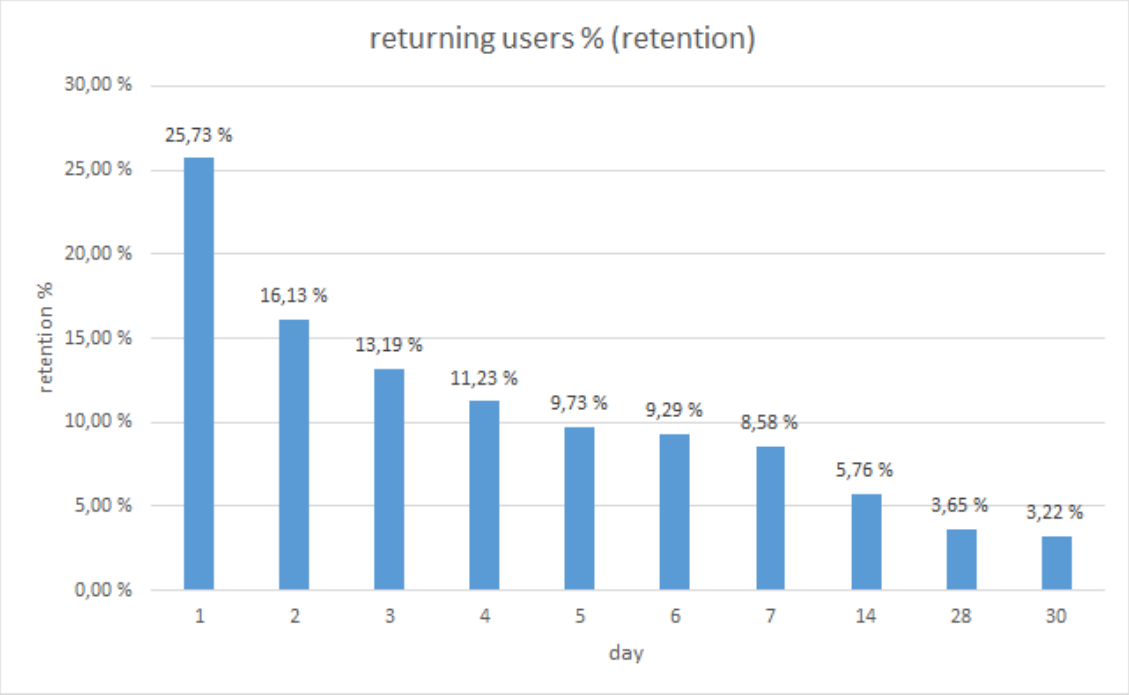

Figure 3. The retention rate measured through 30 days of play.

These numbers show that the retention rate (the measurement of users returning to the game) gets drastically lower as time goes on. Retention was measured throughout the whole testing process (from january to the end of march in 2014) and it shows that in the first month after starting the game for the first time, only 3,22% of the players that originally started the game came back to playing it, with only 16,13% returning to the game after the first day of playing it. This gives some indication to the fact that long term engagement for the players does not exist in the game in its current form and that there is a definite need for things to change, in order for this game to stand as a beneficial product for the development team.

Getting in touch with the users at the start of this 'decline of engagement' can prove to be crucial in pointing out what needs to be improved in the game in order to make it more engaging for the players. As these previous statistics show, players gradually lose interest in the game, which would indicate that this game might not survive at the harsh mobile game market of today.

Testing activities were conducted to get in close contact with players, the potential customers, in order to find out what was causing this issue with getting players engaged with the game. The testing would have a focus on the more softer aspects, like the user experience and the game mechanics, instead of the technical solutions. This was the developer's choice, which supports the earlier claim by Kasurinen & Smolander (2014), that game developers tend to focus on these "more softer aspects" of testing.

3.3 Research Method

The research method used for this thesis was the case study method. A case study method should be used when the type of research question is “how” or “why”, when the investigator has only little or no possibility to control the events under research and the phenomenon under study is a contemporary phenomenon in a real-life context (Yin, 2013). This research is based on a real product currently in development in a game development company. The research was conducted in close cooperation with the company, to better understand their problem and in turn get results that could provide beneficial to their development process. The research question aims to answer the issue of why the players seem to lose interest in Casters of Kalderon as time goes on (as indicated by the previous testing results), and how this issue can be answered through conducting testing.

The research has a focus on a phenomenon in a real-life context, as the game is tested with potential future customers to see what they think about the game, and how that might influence their engagement with it. Case studies can be explanatory, exploratory or descriptive, where the methods in use can be qualitative, quantitative, or both (Yin, 2013). This research is exploratory in nature, as the cause of the issue is not well known, and the data is gathered through questionnaires, interviews and observations, making it qualitative. The research was conducted in one company, with one case, Casters of Kalderon.

In order to get the data to help answer the issue of measuring player engagement, Casters of Kalderon was tested through conducting usability testing with a group of participants. The game was tested on PC platform, but the developers saw that the issues that would be found could be directly applied to mobile platforms also, as the builds are in no way different. The actual testing was done both remotely and on-site in a testing lab, in order to get good coverage and more participants. The developers of the game also provided previous testing results for analysis to help support the testing activities. The main point of the testing was to find out how the experience is for the players. Do they feel immersed in the game and find the goals and the fun factor in it, or do they find it dull, confusing and frustrating with no intention to return to it again later. The developers of the game were very cooperative and regular meetings about the progress of the research were held almost weekly, where testing plans were made, results would be gone through and future actions would be planned for.

Initially there were 10 participants willing to take part in the testing activities, 3 of which were invited to attend the testing lab, and the remaining 7 doing the test remotely and reporting their findings through the questionnaire form. All of the participants were male, ranging from 26 - 31 years in age, and all having previous experience with gaming. The testing lab activity was conducted at the University of Oulu, where 3 participants tested the game on a windows PC under observation in two different sessions. One camera was setup to record the player’s face and speech, with another recording software capturing the player’s actions on the screen. In addition to this, observation was conducted on the player’s actions, and they were encouraged to speak their mind during the testing in order to get a better understanding of what it was they were thinking while playing the game. The video and audio were then analyzed together with the observed notes and feedback questionnaires to find out what the players’ feelings were regarding the game. In the end, from the 7 participants that volunteered for the remote testing activity, only 4 took part in the testing activity due to time and resource constraints on their end, this would mean that altogether there were 7 testers, 3 doing the testing on-site and 4 remotely on their own computers.

4. RESULTS

In this chapter the results of all the testing activities are gone through and the problems with the game are categorized and prioritized. Limitations and benefits are also discussed to reflect on the research method feasibility.

4.1 Testing session 1 in Lab

Three testers were invited to participate in the observed testing done in the usability lab of the department of information processing science in the University of Oulu. The participants playing was recorded on video through a screen capturing software and they were also observed while taking notes on their comments and behavior. All of the testers were male, ranging from 25 - 27 years in age and they all had previous experience with playing games on both mobile platforms and computers. 1 hour and 20 minutes of video material was gathered during this first testing session. Each of the participants had half-an-hour to play the game as they wished, with minimal introduction and instructions. After the testing was completed, the participants also filled out a questionnaire with the following questions:

In your opinion, what was the goal of the game?
Please describe how these goals are accomplished.
What aspects did you like in the game?
What aspects did you dislike in the game?
If you could change something in the game, what would it be?
Please describe any problem situations with the game.
How did you solve these problem situations?
What did you find frustrating in the game?
What did you find confusing in the game?
Your opinion on the combat aspect of the game.
Your opinion on the building aspect of the game.
Your opinion on the in-game coalition system?
Your opinion on the spying aspect of the game?
Open feedback.

The testers were also observed throughout the testing process and any observations made were then reported. Over 40 of different issues were noted, which would then need to be prioritized and discussed with the game's developers. Another test session was conducted later, to let the same players try the game again, this time with a set of goals they needed to complete. The players were able to complete these goals, but the same issues that surfaced in the previous testing session, also surfaced here.

The goal of the game was seen as building a city and protecting it from attacks while also gathering resources to maintain the city. To this, the developers agreed as it had been the designed goal for the game. In order to accomplish this goal, the players felt that they needed to construct buildings in order to gather the needed resources and building an army to defend the city and also wage war on other cities. Securing a good economy was also seen as a way to achieve the goals in the game. One tester also thought that the resources are collected through completing missions, although this was not the case. This would imply that there is a slight communication problem when it comes to accomplishing the goal of the game.

When asked about the likeable aspects of the game, the testers listed things like the interesting environment, versatile features, the multiplayer aspect, the graphics, clear goals and a good flow. When asked about the unlikeable aspects of the game, the testers listed things like confusing user interface, insufficient instructions and some aspects of the game requiring huge amounts of patience. The testers also felt that too many aspects of the game were introduced all at once, which resulted in confusion and frustration. Some usability issues were also reported, more on which later.

Problem situations with the game included user interface problems, control issues and missing some important aspects of the game due to the presentation. There was also some confusion with the terminology and language. When asked about how the players resolved these problems, all of them reported that it was through trial and error.

Frustrating aspects included constructing buildings one at a time, unclear instructions and user interface issues, like finding the right buttons for certain functionality and telling the different buildings in the game apart from one another. One tester reported that he was bored of the constant building work in less than half-an-hour, and would like more active player involvement. Confusing aspects included many questions that were not answered in the game, like are construction workers required in order to build and is it possible to build more than one building at once. One tester reported that he felt like almost all of the meanings of the game were unclear. Other confusing aspects included the resources and attacking other players. This confusion with the game mechanics speaks volumes on implementation of the tutorials in the game, and that they do not currently provide sufficient data for the player to get a clear understanding on the game's mechanics.

The construction of buildings was found logical, even though the time it took to upgrade the buildings was felt to rise a bit too steeply. The resource income was wished to be presented more clearly to the player and it was also wished that the placement of buildings would matter. It should also be possible to cancel the construction of a building at any time, which was not currently possible.

The coalitions play a big part in the multiplayer aspect of the game. They were seen as a good thing as an idea, but they should not be mandatory in order to succeed in the game, as one of the testers felt like he prefers more to play solo. The scouting of other villages was seen as a fun addition, even though the world map where the other player's' villages are shown, was found to be humongous and easy to get lost into. At this point the testing

time for the first test session started to end, and none of the testers had the time to try out the combat in the game.

When asked for open feedback, all of the testers said that they would like to return to the game, if some improvements were made. The game was found confusing and there were clearly problems with the user interface. The testers also felt like the introduction to the game's multiple mechanics should be handled more slowly, in order to quickly understand and get in terms with them. It was noted that introducing new aspects of the game slowly and one at a time, would help boost the enjoyment. Two testers of the game compared it to another mobile game called Clash of Clans, and they felt that that game handled the introduction and instructions better than Casters of Kalderon.

Figure 4. The instructions given to the player at the start of the game.

Observations were also made while doing the testing. It was noted that while the game attempts to help the player to get familiar with the game's mechanics through textual instructions, all of the players felt that there was too much of this text, and that too many instructions were given all at once. Some of the players clearly did not bother with reading all of the text provided. This would indicate that the instruction of the game mechanics should be done a little at a time, in order to slowly introduce the player into a yet unfamiliar new game. This supports the earlier claim made by Febretti & Garzotto (2009), that the developers need to make sure that the players are able to reach the fun of the game as easy and quickly as possible.

A lot of confusion also stemmed from the user interface with its many buttons, that were felt to not represent the actions that they do, and what the player can do with them clearly enough. This suggests that the actions the different buttons have need to be represented in the graphics. There were also barriers in the language use, where the amount of time the user has to complete a mission, was interpreted to be the time that it actually takes for

the player to complete the mission. This would indicate that the use of the English language in the game also needed some improvements.

Figure 5. Buttons of the user interface that the players found confusing

Another user interface issue came in the form of notifying the user of the happenings in the game. The notifications are done in the form of messages from the game, informing them when they are progressing in the game. The notifications are introduced in the form of a red circle, with a numerical value. A brief sound effect also plays in an attempt to get the player's attention.

Figure 6. The notification button

Many of the testers heard the sound, but were not able to deduce that it came from this notification, and that the game had something to tell them. This meant that the testers would continue playing the game, not knowing that they had access to important information in the game. The biggest problem this introduced was that one of the testers was not able to continue the game and got stuck, since in order to continue playing, he needed to check the notification, and receive important information that would help him continue (see **Figure 7**). This clearly caused frustration and confusion in the tester.

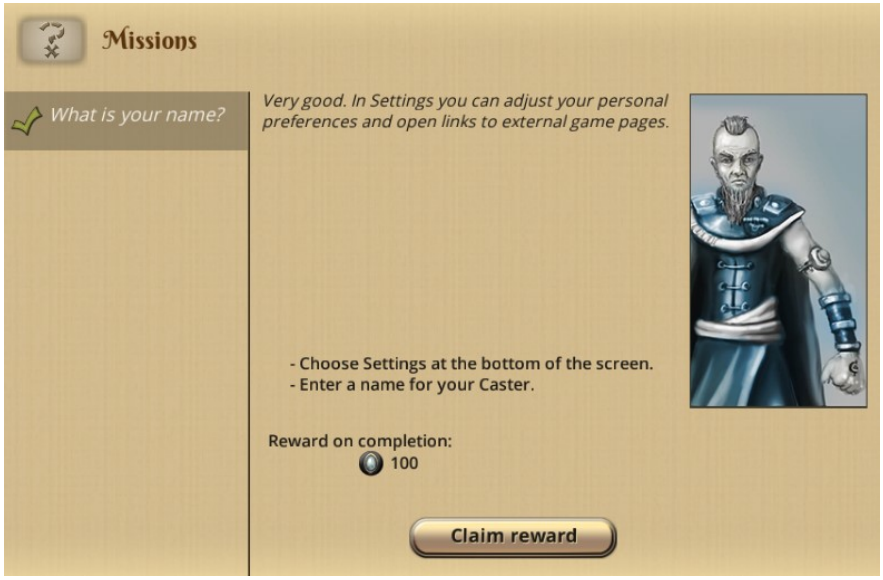

Figure 7. The important information behind the notification, missed by one of the testers.

More elements to indicate the game's progress were also felt as required. The amount of buildings the player has, which of the buildings the player already has built, and on what "level" the buildings currently are, where all information that can be found in the game, but the players had difficulties in doing so. This would indicate that in order to decrease confusion, more of these signs of progress need to be made clearer for the player to see. Due to the different buildings looking so similar in appearance (see **Figure 8**), there was a lot of confusion about which of the buildings serves which function in the game, like the farms that provide food for the player's troops, and the barracks where the player is able to recruit more of these troops. This would indicate that there needs to be a clear distinction between different buildings and the functions that they provide.

Figure 8. Different buildings with their own functionality in the game.

4.2 Testing session 2 in Lab

Another testing session was conducted with the same participants. This time the players were given a set of tasks they were to complete, instead of the free hands on approach of the previous test session. The tasks included:

Build the building "marketplace".
Put a resource of your choice for sell in the marketplace, at the price of your choosing.
Build the buildings "Swords Arena" and "Barracks".
Level up both of the buildings to level 3.
Hire 8 Calisian Sword units.
Hire a hero character from the tavern.
Find a "witch's hut" and a "bandit camp" from the world map.
Send 4 Calisian Swords to raid the witch's hut, and another 4 to raid the bandit camp.
Check the results of the raids.
Find another player's village from the world map, and scout it.
If there are soldiers in this village, send your army to attack it.
Check the battle report to see the results of your attack

It is possible to see this battle in 3D, try to find this feature.
Check to see if you have created a coalition. If not, create one.
Invite other players into your coalition.
Write a message to your coalition's message board.
Send a message to another player.

The results of this testing brought up new issues in addition to the ones already discovered in the previous testing session. While the building portion of the game was already familiar to the player and the players were successful in completing these tasks, problems started to occur when the players were introduced with new game mechanics, like the scouting and battling.

The scouting in Casters of Kalderon is handled by the player controlling a bird like creature that flies over the other player's village. The bird must fly through a set of blue rings in order for the scouting to be successful. The main issue here was that both the purpose of the scouting and the controls of it were not properly introduced to the player.

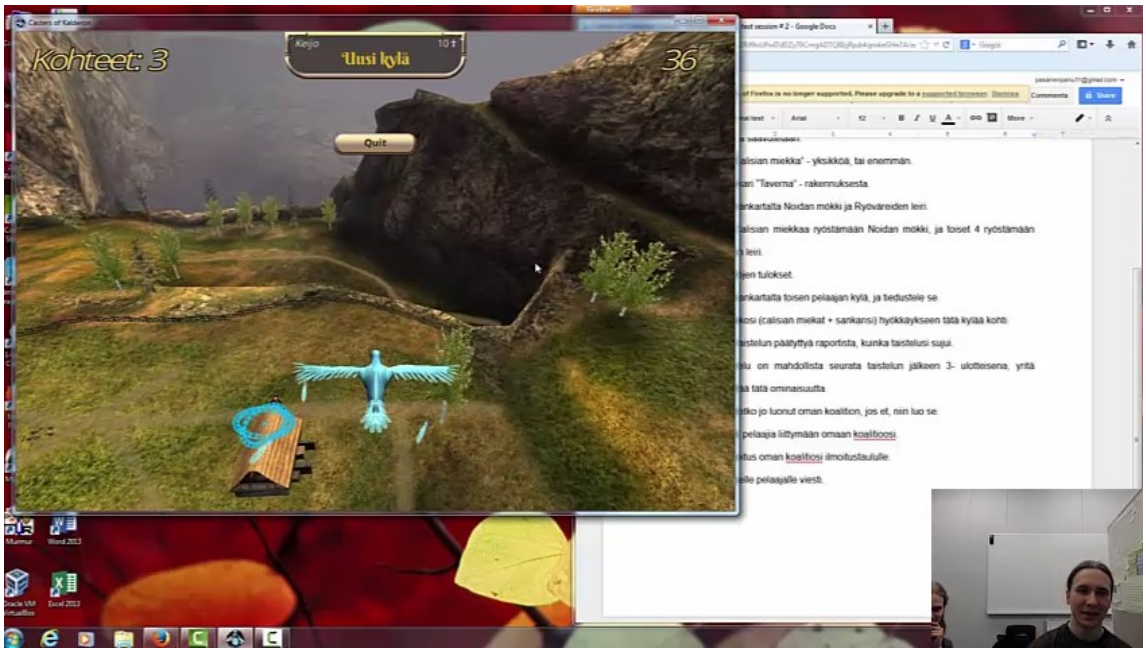

Figure 9. Recorded testing activity with the scouting mechanic.

What happens here, is that the bird starts flying and the player is instantly given control to it, without any information given to the player that he has now assumed control of this new character in the game. This resulted in the bird flying through the air, while the players watched it, unaware that they actually had control over its actions. There are also numbers and text elements introduced to the player, with no explanation as to what they mean. The players wished that there was a simple tutorial screen explaining the controls of the bird, and the purpose of scouting in this context. As the flying was found to be a fun mechanic, once the players were able to get into it, a brief tutorial and additional explanation would ease the accessibility of new players when it comes to this mechanic.

This insufficient information flow to the player continues to the world map. As the players were asked to find a witch's hut and a bandit camp from the map, there were difficulties in telling these objects apart from the map itself (see image 5.) This issue could be avoided if the players were given a visual cue as to what these objects in the world map actually are, like textual descriptions or representative icons.

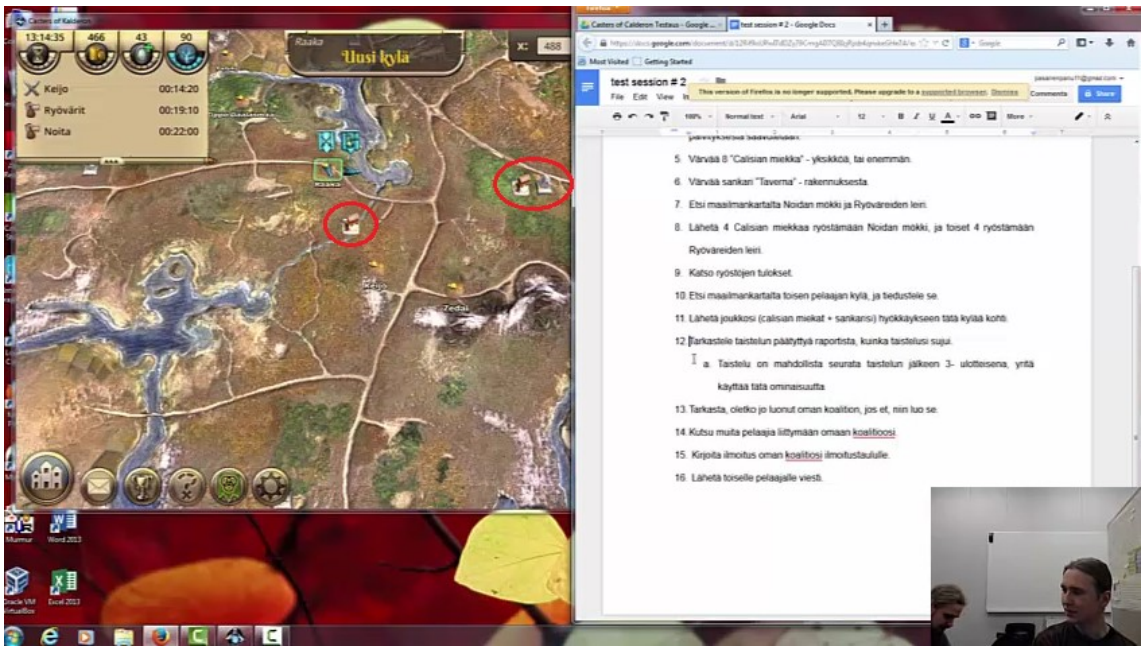

Figure 10. Witch's hut and bandit camp on the world map (circled in red.)

The battle mechanics of the game introduced new issues. The player can send armies to fight other players on the world map. The issues with this aspect indicated lack of meaningful feedback and impact of battles, as players were unsure when the battles were actually taking place and how well their army managed to fight. The testers were also unable to tell that they had actually been attacked by another player, which would indicate that the information flow to the player when it comes to the battles is lacking. By increasing awareness on the situation of the battle, players could be able to feel tension and actually start caring for their army and the outcome of a battle situation. Additional graphical and sound elements could help to increase player awareness in this case.

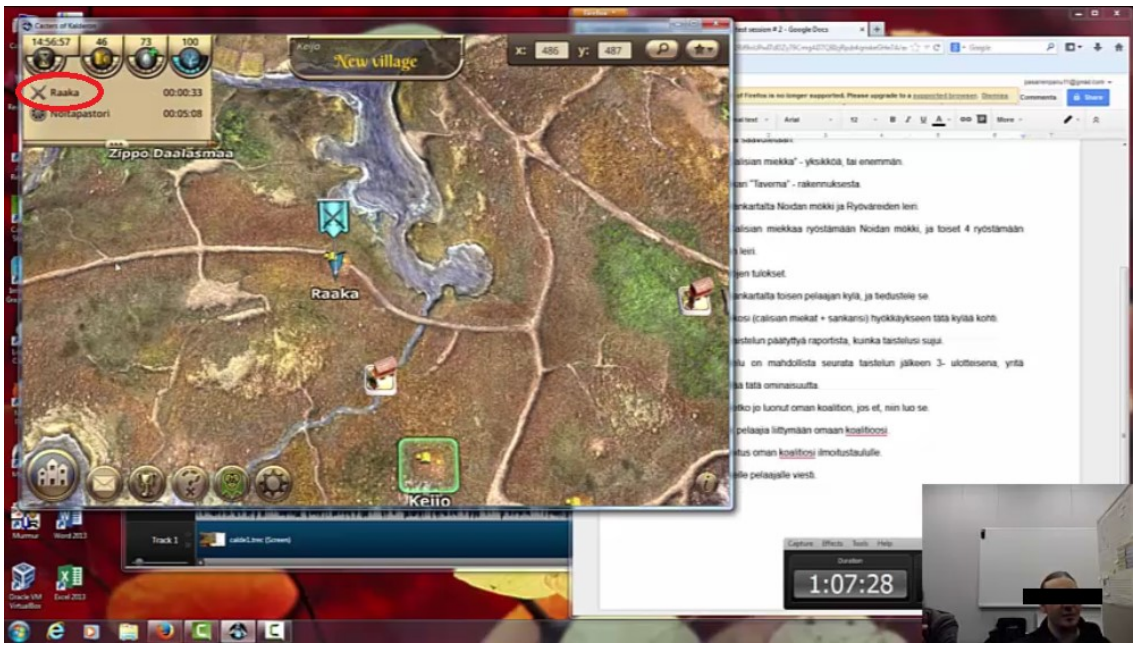

Figure 11. All the information the player gets when a battle is going on as shown by the red circle.

4.3 Remote testing

The testers doing the testing remotely were given a time span of 1 week to do the testing on their own time. A total of 4 testers took part in this testing and the results were gathered using the same questions as in the first on-site testing through a google form, as well as recorded video footage provided by two of the participants.

When asked about the goals of the game and how to achieve them, the answers were largely the same as on the on-site testing. The goal was seen as building the player's village, recruiting an army and both attacking other villages, and defending the player's village from oncoming attacks. One participant even had the idea that he had the ability to develop his player character, which was not the case however.

Liked things about the game included the fantasy setting, the graphics, as well as the building, and scouting aspects of the game. Disliked aspects were that the game was seen as too slow, and the testers felt that there was a lack of meaningful things to do in the game. It was wished that the player character had more deeper role in the game. One tester even said that after a couple of clicks, he felt that there was nothing else to do, but wait for the player's buildings to be completed. It was wished that the game would be a lot faster and easier to approach. More player participation was wished, as most of the actions in the game are based on waiting their completion. Some issues with the UI were also noted, which in turned caused confusion in the testers. Important information in the game was not presented clearly for the player, and they had to search for the information they felt should be provided for them automatically.

One tester noted in the open feedback, that he felt that his village lacked personality. Something that would make the village feel like it belonged to him, and would tell it apart from the other player's villages. This made the tester feel that he didn't feel the need to return to the game to see how his village was doing, as it felt that it was not his own. This would imply that some players feel the need for a sort of personification in games they

play. Something that tells the rest of the game world that this is “them” in the game. An interesting find that i think could be researched further.

In conclusion these remote testing activities were difficult to manage. Players were given clear objective and a timeframe for the testing, but some of the already volunteered testers dropped out after the deadline had already finished. The resulting questionnaires gave similar results to the testing done in the lab, which was a good outcome, but this remote setup made the results harder to communicate, as all of the feedback was only in textual format. I would argue that this kind of testing is best done with an observer on site, in order to get the communication of the issues with the game better.

4.4 Issue prioritization

Once the testings had been conducted, the resulting questionnaires, video materials and observed notes were delivered to the developers, who were really positively surprised by the results. Even though had had plenty of experience with the game themselves already, the testing was able to bring up new aspects that they had not thinked about at all.

The developers wanted to prioritize the resulting change proposals that were provided into three categories, low, medium, and high with high priority issues being the most important for changing to game more to the liking of the testers. Major problems were seen as the issues that were brought up by multiple testers, and in multiple test sessions, and issues that might hinder the player progress in the game.

Most of the high priority issues had to do with the tutorials and introducing the player to the game’s mechanics, as well as some issues with the user interface, like the lack of icons for different buildings to make them easier to identify. Medium priority issues included small tweaks to the user interface, like additional textual elements and better introduction to some of the game’s aspects. Low priority issues included issues with the gameplay as well as suggestions for new elements to increase the enjoyability and immersion of the game, like the personalization and enhancement of the role of magic to increase player immersion, but these issues would require some heavy redesign of the game and are most likely not implemented.

The developers found the user feedback very important, and showed clear willingness to fix the issues that came up during testing. The testing work clearly focused on the player experience and the game mechanics, and customer-based-feedback had a high emphasis, which supports earlier research conducted by Kasurinen & Smolander (2014).

The categorized and prioritized testing results are going to be used to tune the final product to better cater to the intended customer audience as closely as possible, which also supports this earlier research. Prioritizing the issues was found helpful, and the development team gathered to discuss which changes would be made and at what amount of effort, at a later date.

The issue list can be found in **Appendix A.**

4.5 Limitations & Benefits

There were a few limitations to the methods that were used making this thesis. As this testing was conducted by only one person, the observations are based only on that one person's notes and findings. With more testers, more detailed results could be gained from these activities. As pointed out in previous research by Niels Jacobsen & John (1998), no single evaluator will detect all the problems when analyzing test session. Another limitation was that the testing participants were all male, 25 - 31 years in age, with previous experience with games of different genres. The results were very similar, but they are not easy to generalize into other audiences, like children and teenagers, players of the other gender or people with little to no previous experience with games. It was a request by the developers, that the demographic should not play a big part in the testing activities.

Another issue was the amount of participants willing to take part in the testing. From the initial 10 participants that signed up, only 7 took part in the testing, with 3 participants doing the testing on-site at the testing lab, and the remaining 4 doing the testing remotely on their own computers. More participants would have yielded more results for analysis, which could have provided better data, but even with this amount of participants, the developers were able to gain meaningful insight on the aspects of the game that need improving, and were willing to act on the feedback gained from them.

The additional video footage that was provided by the remote testing participants did not give any meaningful data on the issues, as there was no sound available, and no video recording of their faces. From this it can be concluded that in order for these kind of testing activities to be feasible, a third party observer should be at the scene of the testing taking notes and discussing with the participants about their feelings with the game. The video material recorded with sound in the testing lab was really beneficial, as the players spoke out many of their thoughts while playing the game, which could then be written down. The recordings were also beneficial in the way that the only observer that was there was very busy with both observing, discussing and taking notes while the testers played the game. This allowed for many additional notes to be made after the testing was concluded. In total 3 hours and 10 minutes of video material was recorded for analysis during the testing activities done in the lab.

5. CONCLUDING DISCUSSION

The method of getting the users in the same space with the observer for the testing was a really good way to gain feedback. While usability methods were applied, like observing the players playing the game, giving the users tasks to complete and a questionnaire to give feedback on, the methods were applied loosely, in order to respond to the fact that the product under testing was a game, which gives the player a lot to do with and many functions to experiment and explore. Restricting the testing too much might not have given such good results.

While the testers found positive aspects in the game, like the fantastical environment, the graphics and the multiplayer features, most of the feedback brought up things that need improvement in the game, in order to keep the players interested in it and returning back to it.

Confusion with the game mostly stemmed from the fact that the players are not given clear enough goals to aim for at the start of the game. Most of the games functionalities are introduced in the form of a massive textual introduction explaining what the different elements of the user interface can do, causing an overflow of information requiring too much attention and memorization from the players. Most testers would have preferred if the game introduced important aspects of the game, like the building construction, scouting enemies and the battle elements, little at a time, allowing for new players to get familiar with the game's many mechanics. As Andersen et. al (2012) pointed out, it is important to design early levels of game in a way that enables the players to experiment and discover these game mechanics.

The game has a set of goals in form of tutorials, to help the players get familiar with the game, but these are not properly introduced, as the players are able to do things outside of these tutorial goals, that provide nothing when it comes to accomplishing them. The player's progress on completing these goals are also not well informed for the player, causing even more confusion. Many of the players wished that the important functions of the game would be introduced only a little at a time, to gain a better understanding of the game's mechanics, upon which they could later experiment and explore upon, as their mastery with the game increases. The current tutorials are introduced in form of large amounts of detailed texts, which works against players learning more from exploration than from reading text, as pointed out by Andersen et. al (2012).

User interface problems also posed a barrier in accessibility, when the players were not certain what functions are hidden behind which element, like the buttons in the interface. The functionality of the different buttons are introduced at the very beginning of the game (see image 2), and can later be accessed through using a help button. In previous research on game tutorials conducted by Andersen, et. al (2012), they found little evidence that adding a help button into a game helps player progress, and it was actually found to reduce player progress by 12%. From this it can be concluded that throwing too much information at the player from the get-go is not a good way to get them to understand the point behind it, even if that information can be accessed later. The introduction as it is now, would require large amounts of rework, but it would definitely pay off in the long term.

One major usability factor was the difficulty in telling the buildings that the player is able to construct apart from each other. With additional graphical interface elements, this problem could be solved. The game could have the building names available at all times, so the players could identify them based on their name. This graphical element could also provide additional information, like the building level, in order to keep the player informed of their progress at all times. As the buildings are now, they resemble each other too closely for the player to easily differentiate them. Use of simple representative icons for the buildings could also help alleviate this issue.

Even though problems understanding the user interface never stopped the testers from playing the game, it introduced some barriers in their progress. Even though the players were not always sure what a certain button in the game would do, they would use trial and error to find out the functionality, which was found both confusing and slightly irritating. This supports earlier research by Febretti & Garzotto (2009), which stated that players tend to find strategies to overcome usability problems that occur during gameplay. This would suggest that the game needs to improve the graphical elements, like the buttons, to better represent the actions that they do, and help the players to get into the fun of the game more easily, in order to avoid the temporary decrease in engagement (Febretti & Garzotto, 2009). Even though the players were willing to sit through their frustration and continue playing, it cannot be said that it was not only because of the testing environment and setting. What about the players who buy the game on their own volition, and then find out that the game can be confusing and frustrating? These issues with usability and the quality of user interface play a very important role in the player experience, and they play a major part when the customer makes a decision whether or not to actually pay for a game (Rajanen & Marghescu 2006).

When it comes to getting the players engaged with the game, there were few factors that hindered this. When asked about if the testers would see themselves returning to the game, they were willing to do so, if some improvements were made. The testers felt that the flow (Chen, 2007) of the game needed some improvement, which refers to the level of immersion discussed earlier and supported by Brown & Cairns (2004). For example the introduction of the mechanics being done only a little at a time to provide a smoother flow into the game. This supports the earlier claims by Cowley et al (2008) that the game mechanics should start out simpler, and then advance in complexity as familiarity and mastery with the game increases to support exploration and experimentation. Exploring and experimenting were also defined as activities the player performs to increase the desire to continue playing (Schoenau-Fog, 2011).

The issue of lack of personality was also brought up by one of the testers. The tester felt like by introducing something that enables them to give their town a personality, they would be more likely to return to it more regularly. This could indicate that the level of immersion, and thus the level of player engagement with the game could improve, if the players could bring an element of themselves into the game. However this point was only brought up by one of the testers, so more rigorous testing on this aspect is required to see if it brings added value to this case. This would be an interesting topi for research in itself.

In conclusion it can be said that Casters of Kalderon as it is now, has both positive and negative aspects, with the negative aspects having a strong effect on player enjoyment and through it on engagement. The tutorials as they are now do not provide enough information for new players to get in touch with the game's many mechanics, and should be modified to give the player the necessary information as it is required. The players could then build their experience and mastery upon these newly learned functionalities,

and then be introduced to new ones as the game is progressing, which again supports the claim made by Cowley et al (2008).

To improve the game in the long run, more of similar testing activities should be done in order to make sure that the changes have had an effect on the player's opinions on the game. These testing activities should be done iteratively after every change, and it is of utmost importance to include an observer into these testing activities, as that was seen as the most efficient way to gather information from the players. It would seem that complex games like this require a strong quality assessment process (Febretti & Garzotto, 2009), which Casters of Kalderon lacked, since the development company did not have any kind of existing QA process.The same could also be said for other games of the same genre, as pointed out Andersen et. al (2012) in their previous research, as tutorial effectiveness depends on the complexity of the game. The more complex the game, the more of these testing activities need to be conducted in order to make sure that the players, and the potential customers, are able to grasp the concept of the game at the earliest possible state.

As tutorials were the key point for introducing players to the new game, for future research it would be good to test tutorials on their own. Going through a tutorial with the testers from start to finish could have given better information on what the actual factors were that made the players like or dislike the tutorial. Finding out that the tutorials were too information heavy in their current state was very beneficial for the developers, but actual instructions on how to improve them significantly could not be deduced from the testing. One major point was that the tutorials need to give the players small pieces of information one at a time, to slowly introduce the player into the game, and make it easier for them to get a grasp of its mechanics. This would provide them with a smoother flow (Chen, 2007, Sweetser & Wyeth, 2005) to get involved and invested in the game. It is good to keep in mind that these results apply best for games of the same genre, strategy games, as they are often complex games with many functionalities for the players to get familiar with. For smaller, more quicker to learn games, like arcade games, the notes gained from this testing might not apply.

The developers of Casters of Kalderon found these testing activities to be very beneficial. The importance of getting actual players involved was very clear, as was the importance of having a person observing the game play situation and gathering feedback both during, and after testing. The issues found during testing were categorized and prioritized and they will be discussed by the development team and relevant changes to the game will be made at a later date. Usability testing was found to be an effective way to find the issues in the game. Even though developers might have a clear understanding on how a game is supposed to work and what the different mechanics are behind it, these things are not easy to communicate to new players who are unfamiliar with the game. Tutorials play an important part in introducing these game mechanics, and they need to be designed and implemented in a way that helps the player to get comfortable with the game's many mechanics as easy and quickly as possible. The issues found in Casters of Kalderon can very well reflect to other games of the same genre, and this thesis provides good insight on what issues need to be kept in mind when designing similar games, or complex games in general. This can help both future game development and research on the subject, as it provides insight on an actual case, what went wrong and what needed to be improved in order to make the experience better.

References

Andersen, E., O'Rourke, E., Liu, Y., Snider, R., Lowdermilk, J., Truong, D., Cooper, S. & Popovic, Z. (2012). "The Impact of Tutorials on Games of Varying Complexity". CHI'12.

Brown, E. & Cairns, P. (2004). "A Grounded Investigation of Game Immersion". University College London Interaction Centre. CHI '04.

Chen, J. (2007). Flow in Games (and Everything Else). Communications of the ACM, April 2007, Voume 50.

Cowley, B., Charles, D., Black, M. & Hickey, R. (2008). "Towards an Understanding of Flow in Video Games". Computers in Entertainment. Volume 6 Issue 2.

Desurvire, H., & Wiberg, C. (2009). Game usability heuristics (PLAY) for evaluating and designing better games: the next iteration. In *Online communities and social computing* (pp. 557-566). Springer Berlin Heidelberg.

Desurvire, H., Caplan, M., & Toth, J. A. (2004). Using heuristics to evaluate the playability of games. In *CHI'04 extended abstracts on Human factors in computing systems* (pp. 1509-1512). ACM.

Dumas, J. S. & Redish, J. C. (2009). "A Practical Guide to Usability Testing". Intellect Books. 1999.

Febretti, A. & Garzotto, F. (2009). "Usability, Playability, and Long-Term Engagement in Computer Games". University of Illinois & Politecnico di Milano. CHI EA '09.

Federoff, M. A. (2002). *Heuristics and usability guidelines for the creation and evaluation of fun in video games* (Doctoral dissertation, Indiana University).

Järvinen, A. & Mäyrä, F. (2002). Communication and Community in Digital Entertainment Services. University of Tampere. Hypermedia Laboratory Net Series 2.

Kasurinen, J. & Smolander, K. (2014). "What do Game Developers Test in Their Products?". Lappeenranta University of Technology. ESEM' 14.

Marghescu, D. (2009). Usability evaluation of information systems: A review of five international standards. In Information Systems Development (pp. 131-142). Springer US.

Nacke, L., Schild, J. & Niesenhaus, J. (2010). "Gameplay experience testing with playability and usability surveys - An experimental pilot study". Playability and player experience: Proceedings of the Fun and Games 2010 Workshop.

Nacke, L. (2009). "From Playability to a Hierarchical Game Usability Model". Proceedings of the 2009 Conference on Future Play.

Nacke, L. & Lindley, G. (2008). "Flow and Immersion in First-Person Shooters: Measuring the player's gameplay experience". Proceedings of the 2008 Conference on Future Play.

Pinelle, D., Wong, N., Stach, T., & Gutwin, C. (2009). Usability heuristics for networked multiplayer games. In *Proceedings of the ACM 2009 international conference on Supporting group work* (pp. 169-178). ACM.

Pinelle, D., Wong, N., & Stach, T. (2008). Heuristic evaluation for games: usability principles for video game design. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1453-1462). ACM.

Rajanen, M., & Marghescu, D. (2006). The impact of game usability to player attitude. In Proceedings of 29th Information Systems Research Seminar In Scandinavia (IRIS 2006), Helsingoer, Denmark.

Rajanen, M. & Nissinen, J. (2015). "A Survey of Game Usability Practices in Northern European Game Companies". University of Oulu. IRIS: Selected Papers of the Information Systems Research Seminar in Scandinavia. Issue 6 Paper 8.

Sweetser, P. & Wyeth, P. (2005). "GameFlow: A Model for Evaluating Player Enjoyment in Games". Computers in Entertainment. Volume 3 Issue 3.

Schoenau-Fog, E. (2011). "The Player Engagement Process – An Exploration of Continuation Desire in Digital Games". Proceedings of DiGRA 2011 Conference: Think Design Play.

Yin, Robert K. (2013). "Case Study Research. Design and Methods". Sage Publications.

Appendix A. Issue list

Please note that the final issue list had colors representing issue priority. This was found to be very helpful in the categorization of the issues. Due to this thesis being a print, these colors are not shown.

Category	Issue	Priority	Also in test # 2
Tutorial	Remove the huge textual introduction in the start of the game.	high	
Tutorial	Tutorial rework / Better introduction.	high	x
Tutorial	Tutorial to world map / battles / scouting.	high	x
Tutorial	Show the player the progress of tutorials (1/2 farms built etc.).	med	
Tutorial	Moonstone explanation (what they are, what can be done with them, how to get more).	med	
Building	Notifications for completing construction of a building	high	x
Building	Unlocking buildings as the game progresses.	high	
Building	Add functionality for cancelling buildings in progress.	med	
Building	Make the benefits of upgrading a building more clear. (A stat bar of sorts?)	low	x
Building	Allow the players to build multiple buildings at once (more queues from more builder's hut).	low	
Scouting	A quick tutorial for the controls and the goal of scouting with the bird.	high	x
Scouting	Indication that the running number is actually a timer.	low	
Battle	Show the battle replay by default, when the player checks the battle report.	high	x
Battle	Enhance the feedback of battles through notifications, sounds and extra graphics.	med	

Battle	The difference between raiding and attacking? Raiding seems obsolete at the moment.	low	
Battle	"Confirm" or "To battle" button to help the player confirm he is ready to battle.	low	
Battle	Better indication when a player is under attack / attacking another player.	low	x
Coalition	Notification on the state of inviting another player to join the player's coalition.	high	x
Coalition	Button for accepting a coalition invitation (appears automatic)	high	
Coalition	Improve player awareness of messages sent by coalitions / other players	low	
UI	Better notifications (too easy to miss as they are now). Change the graphic/sound.	high	x
UI	Make the buttons of the UI more representative of the function they have.	high	x
UI	Building names visible or Representative icons for buildings	high	x
UI	Landmark names/icons on world map (witches hut, bandit camp, etc.).	med	x
UI	Improve coalition message system.	low	
UI	Maximum amount of resources presented as a numerical value, instead of a progress bar.	low	
UI	Some of the heading banners (building names etc.) closely resemble the button elements.	low	x
UI	Change the term "village upkeep" to "food", in the village view.	low	x
UI	Language selection to the start screen.	low	
UI	Building levels visible.	low	x
UI	List of buildings in build menu to a grid.	low	x
UI	Make the battle replays more accessible.	low	

UI	Move the hero characters to the center with the other units in battle preparation.	low	x
Gameplay	Benefits for building placement?	low	
Gameplay	The role of magic could be enhanced. More spells & benefits from using them.	low	
Gameplay	More functionality for the third-person-perspective. Doesn't serve a purpose at the moment.	low	
Gameplay	Provide a way for players to make their village more personal (banner flag, logo, etc?)	low	