Gamification in Mobile Language Learning: Improving User Satisfaction for Norwegian Immigrants

Morten Grina Myhre

Master thesis

Master in Interaction Design

30 ECTS

Department of Computer Science and Media Technology

Gjøvik University College, 2015

Avdeling for informatikk og medieteknikk
Høgskolen i Gjøvik
Postboks 191
2802 Gjøvik

Department of Computer Science and Media Technology
Gjøvik University College
Box 191 N-2802
Gjøvik Norway

Abstract

During the previous years, technology has been used for assisting immigrants with language learning in Norway. Using technology to teach languages contributes to increased effectiveness in learning for students. Students perceive this as more fun. Now that mobile devices have been developed to the point that they are powerful mini-computers, the potential for learning with these devices is getting bigger. That means people can learn by using interactive technology whenever and wherever they would like to, using their own personal devices. M-learning is used to define the applications that are used on mobile devices to assist learning. Using m-learning is evidently increasing motivation for learning. To further increase motivation, gamification can be used as a design approach. The target group in focus of this thesis is immigrants in Norway that currently learns the Norwegian language. The thesis focuses on how to implement gamification in mobile language learning applications, and if this affects the students satisfaction during learning. Initially, qualitative user studies were used to gather an understanding of the current situation on e-learning in Norway - how e-learning is currently used, and how receptive the current students were towards a digital alternative to learning. After this information was gathered and a prototype was designed, an experimental design was developed and conducted. Data was gathered to form the hypothesis and result of the thesis. The thesis concluded with that gamification, even though it pointed to increased satisfaction, could not increase the users perceived satisfaction.

Acknowledgements

I would like to thank the following mentions for assistance during this thesis:

Cristina Marco – For assisting me throughout the whole process and telling me that I did not have ample time, even though I felt like it. You have been a very good motivator and help through the project period. Thank you for being my supervisor.

Frode Volden – For helping me analyzing the data carried out by the experiment when I thought I had hit a wall.

Miriam Begnum – For being available whenever there was a problem and always providing detailed and thoughtful help.

Runhild Grina – For helping me make appointments with employees at Gran Voksenopplæring for both user studies and experiments.

Employees and students at Gran Voksenopplæring and Gjøvik Læringssenter – For providing me with data that assisted me greatly in this thesis. Without you, this would not have been possible.

Table of Contents

Word count of content excluding Appendixes and Bibliography: 16, 606.

Table of Figures

1. Introduction

1.1 Project Description

Mobile phones have increasingly become popular as an affordable and necessary product in the current world. It has become a multi-functional tool that is used for communication, consuming media, browsing the internet, navigating by GPS and more. Compared to desktop computers, mobile technology offers higher availability as they can be accessed anywhere and when the user wants to. As the mobile platform is still being explored by software developers and researchers in terms of potential, new applications are increasingly being published through mobile markets. An area that has recently been generating interest is the use of mobile technology as a tool for learning.

E-learning is the use of electronic educational technology (Clark and Mayer 2011, p.8) and its growth has been described as explosive over the previous years (Holmes and Gardner 2006). *M-learning* is currently being used as a definition for using mobile technology to develop a learning experience (Ozuorcun and Tabak 2012), and can be categorized under e-learning. The term is more simply described as e-learning on mobile devices, but with exploiting the advantages and considering the limitations these devices have. Devices as laptops, PDAs, tablets and mobile phones are categorized under the term, since they are all portable. The most prevalent device would however be todays smartphones, which are used for all purposes, and can be used anywhere at any time. The advantages of m-learning are that learning also can happen anywhere at any time. The weaknesses are that the devices, since they are portable, are smaller in size than stationary devices. That means screen size is smaller and input availability restricted, which makes them less suited for reading on and partially less efficient to interact with. M-learning applications have to consider both weaknesses and strengths of these devices.

The fact that mobile devices are emerging as a new platform used for learning, can be seen on both *Google Play* (Google 2015) and Apple's *Appstore* (Apple 2014) where applications have their own education category. Applications that focus on learning languages have become popular on the mobile market as many offers knowledge for free. Since the use of mobile devices is widespread and available to most people, the learning activity can happen whenever a user has the time to do so. Mobile devices are rapidly increasing as a consumable product, and are therefore also used more frequent in trivial tasks as surfing the internet, playing games and listening to music. With such a frequent form of usage, mobile devices may be relevant for more serious tasks in the future. By uncovering their potential they can serve as multipurpose-devices, even to a bigger degree than today. Mobile language learning is developing rapidly, and research is not published at the same rate as technology is advancing. The area is important to continuously research to be able to keep up with the technology. Stockwell and Sotillo (2011) highlight the importance of such research in their request for research on the area of language learning on mobile devices.

Several applications have been developed for language learning. Among them are the applications reviewed in this thesis: *Duolingo*, *Memrise* and *Babbel*. A study done on the effectiveness of learning using the mobile application *Duolingo*, states that it is more effective than traditional learning (in a classroom with direct instructions and lectures) (Vesselinov and Grego 2012). The study results showed that effectiveness per hour was significantly higher, and that users would need only 26 to 49 hours of using *Duolingo* to cover the material of the first college semester in learning Spanish. Educational applications are usually fun, interactive, and time spent is voluntary by the user.

In Norway there are interactive e-learning applications for language learning that are used as a way of supplementing and even replacing some of the traditional learning lessons. *Migranorsk*[1] and *NorskPluss*[2] are e-learning applications used today by learning centers in Norway. Learning centers is used as a collective term for the public adult education institutions for immigrants in Norway. Learning centers offer teaching of the Norwegian language as well as social studies.

[1] URL: http://www.migranorsk.no/
[2] URL: http://www.kunnskap.no/index.php/pages/19,norskpluss.html

The classes in this learning center is divided by proficiency levels and these levels are defined by the *Common European Framework of Reference for Languages (2001)*. The levels are officially ranked in the following ascending order: ***A1, A2, B1, B2, C1*** and ***C2***. To give a rough summary of these levels, **A** levels are beginner or elementary knowledge of the language, **B** levels are intermediate or just above intermediate and **C** levels are advanced and mastered proficiency of the language. The learning centers contacted during this thesis offers language learning ranging from A0 – an unofficial language level used to describe illiterate people – to and including B1 that is a prerequisite for studying at colleges and universities.

Learning centers in Norway are currently using e-learning that is adapted to their students respective curriculum. *Migranorsk* and *NorskPluss* are examples of such e-learning regarding learning the Norwegian language. E-learning is used together with traditional learning – both in school context and outside as a tool for homework. The current solutions that are used are mainly web-based software on desktop computers. The downside of these e-learning applications are that they appear to be quite outdated in terms of design and do not seem at the first glance to have much focus on usability and user experience. What seems to be efficient in learning contexts is a way of learning that keeps the student motivated. Mobile devices seem to have a solution to a problem like this. The potential for supplementing or even enhancing the learning experience is present.

Long sessions of learning may be better to perform through web applications, textbooks or oral teaching. However, the technology of a mobile device and the fact that they are increasingly prevalent, opens up for a potential to supplement traditional learning. The experience should be fun and provide good usability – the user should be focused on the task at hand, and have an easy user interface that requires minimal effort adjusting to. The aim of this thesis will be to incorporate gamification in such an effective design that can supplement and even enhance the current learning methods for learners of the Norwegian language.

In learning, motivation is a component for good results. *Gamification* is a design approach used to build motivation of users and give them a better user experience. Though the definition of gamification is still not formally decided, some explanations are agreed upon. The most accepted

one is "*Gamification is the use of game elements and game design techniques in non-game contexts*" (Werbach and Hunter 2012). Gamification provides is a more fun and engaging experience for the user. By having a simple interface, an effective execution of the task can be performed when the user feels like it. This is in addition helping the motivation of users, as opposed to being required sitting through hours of lectures and courses.

Mobile devices are a new learning platform, and have for that reason not yet generated a foothold in Norway. Some of the current most popular applications on the Android mobile market – *Duolingo*, *Memrize*, and *Babbel* – are teaching multiple languages. These applications are so intuitive that the users can start learning the language self-efficiently. Recurring languages are Spanish, German, and Portuguese. These applications also contain elements of gamification, which may be a reason they are so popular. If something like these applications could be utilized for learning the Norwegian language, immigrants who would like to learn the language may find it easier to find time for tasks when out of an educational context. To approach just one language would mean to tailor the user interface to this particular language rather than make it globally applicable, as the currently mentioned applications are. The target group is also more specific than the general population. A significant amount of the people in the target group is from developing countries in the world, which means that some may not have much experience using digital technology. However, mobile phones are increasingly prevalent, and can be anticipated being used by this target group. Anticipating the target group is daily users of mobile devices, motivation of continued use is important to keep in mind when designing such a solution.

The current technological advances are bringing forth higher possibilities of learning a language through digital devices, and the "newest" devices would be those who are known as mobile, with an emphasis on tablet computers and smartphones. By being able to use devices that are available at all times, language learners can determine themselves when to learn. Also, by using the mobile phone as a tool for homework, language learners can integrate the solution in a personal device they may be using for other tasks. This means they can use a device they are familiar with, which may improve the effectiveness when used. Increased effectiveness may also lead to better results in terms of learning outcome over time. An effective learning outcome is also heavily based on the users' level of motivation. To develop a user interface for such an

application, the positive sides of mobile learning needs to be analyzed. Further, the technology of today's mobile devices will also be needed to be considered as of how to be utilized to its potential. Screen sizes, input availability and mobile network are some restrictions which will need attention.

This study is relying on using methods from Interaction Design to design a prototype for language learning, as well as using the prototype in an experimental design. The methods will be used for collecting both qualitative and quantitative data by interacting with users of e-learning. The target group of users will be learners of the Norwegian language attending Norwegian learning centers. What may have to be considered with this target group is that most of them have a limited understanding of the Norwegian language. Big parts of the target group are originated from less developed areas in the world, which needs to be taken in consideration. Users seem to have an opinion that interactive learning on mobile devices is more fun and engaging than traditional learning. Users learning in their own pace as well as getting rapid feedback and results are critical factors for the effectiveness of these devices (Trentin and Repetto 2013). To discover the users' needs and wants, initial descriptive research methods will be used to gather important perspectives and thoughts from potential users. A prototype will be based on preferences of the target group, current usage in Norwegian learning centers and by reviewing existing solutions. The final experiment will measure the perceived satisfaction of each user trying out a gamified and non-gamified language learning application. Satisfaction is defined by how pleased the user is with a product or service. According to Calisir and Calisir (2004), satisfaction can be determined by the usefulness and learnability of the product.

Usefulness

Usefulness is a usability goal according to Rubin and Chisnell (2008). It describes to which degree the users can achieve what they want and need, as well as how likely they are to frequently use the product. It is described by Rubin and Chisnell as the goal that is most important to fulfill, since the system can be easy to learn, efficient to use, and effective, but not fulfill the need of the user.

Learnability

Learnability refers to how easy the proposed product is to learn for first-time users, as well as how long the skills are retained by the user (Lazar, Feng et al. 2010). The users should spend a minimal amount of time and effort to learn how the system works. The users will have to be self-reliant during the majority of time spent using the application. Learnability can be perceived by how easy it is for a first-time user to operate a system, or counting the amount of errors a user encounter when performing a task for the first time.

The more satisfied a user is with a product, the more likely the user will return to use it as well. Satisfaction and extrinsic motivation are two closely related terms that can be the solution to a lack of motivation in learning contexts.

1.2 Research questions

This thesis will work with how to develop a gamified user interface, and which gamified mechanics that should be included for a Norwegian language learning application on mobile devices. The thesis will take advantage of methods used in the field of Interaction Design to reach the desired result. It will ultimately contribute with concluding whether or not gamification can improve satisfaction in Norwegian language learning on mobile devices, and if gamification in design can be used in non-gaming systems for an improved user experience. It will also research how gamification should be applied to a graphical user interface in a language learning context. Gamification is still a relatively new area within research, and further research is encouraged to extend the field. The application design process will focus heavily on the user group and current state of the art technology. The design of the user interface will be aimed at people with low technological competence, but should also be interesting for those who are more experienced with the use of technology. The language is also an issue, as users may not have the required language skills to understand the textual content of the application. The target group is international, and universal conventions will have to be used. The research will study habits and preferences of language learners in Norway in regards to current technology. Based on data

gathered through qualitative field studies and experimental research, results will bring forth whether or not gamification can be a factor for satisfaction in mobile language learning. The result will hopefully prove to students, teachers and facilities that m-learning can be a valuable tool for learning in the future.

As a topic, this thesis will focus on gamification mechanics in mobile language learning. The project will use research methods from the area of Interaction Design and Human-Computer Interaction to create knowledge. The research will study how to implement the mechanics and see if these mechanics provide improved satisfaction of the users. The main research question is **"How can *gamification* contribute to increased user satisfaction in an interactive mobile language learning application".**

To answer this question, it will have to be divided into additional sub-questions.

Sub question 1: Which gamification mechanics could be used in mobile language learning applications to improve user satisfaction?
This question addresses which gamification mechanics (points, leaderboards, story, etc.) should be applied to language learning applications based on what is going to be achieved through the application by the users. This will be answered by reviewing state of the art applications, reading earlier studies on the topic and by conducting qualitative user studies at the learning centers.

Sub question 2: How should a mobile language learning application be designed when using gamification as an approach?
This question addresses the design of the application. How would the user interface be designed when a gamification design approach is used? How should the application be designed when thinking of the user group? It will be answered by reviewing state of the art-applications, and by collecting data from qualitative user studies.

Sub question 3: To what extent will a gamified design improve user satisfaction?
Does gamification notably affect user satisfaction in a language learning application? This question contains the independent variable *gamification* and the dependent variable *satisfaction*

that are the main variables of this thesis' problem statement. By doing a research experiment on users with two comparable conditions (One condition with gamification and one without), this research question will be answered.

Shortly summarized these question will be used to answer which mechanics to include, how to design the application by using a gamification approach, and ultimately compare a gamified and non-gamified design to explore if gamification increases user satisfaction.

1.3 Explanation of Terms

Some frequently mentioned terms throughout the thesis is explained in this section.

Ubiquitous: Ever-present or always available.
Gamification: The use of game-related elements in a non-gaming context.
E-learning: Electronically/digitally assisted learning.
M-learning: Mobile learning / Learning performed on mobile devices.
Web application: Software that is developed for using through an Internet browser.
Mobile application: Software that is developed for the using on mobile devices.
User-centered design: Design methods used in Interaction Design. Each method used is explained in details in the Methodology section.
Learning centers: The common term for where immigrants in Norway learn Norwegian and social studies. In this thesis, Gran Voksenopplæring and Gjøvik Læringssenter have been cooperating.
SUS: System Usability Scale. The questionnaire form used as a measuring instrument in this thesis.

1.4 Pre-study

During the autumn semester of 2014, a study was performed by the author on how to develop an interactive learning application for immigrants in Norway learning the Norwegian language (Myhre 2014). The project was carried out in the course “Specialization Course 2”. The result and conclusion of that project is discarded for this thesis, since the direction of the goal had changed. The project was focusing on how to develop a mobile user interface for learners of the Norwegian language by using user-centered methods. This thesis is compared to the previous project aiming to measure the satisfactory outcome of users by including gamification in a mobile language learning application. Even though the aim is not the similar, the qualitative studies conducted in the course contains data that is relevant to this project. The interview results and questionnaire results are therefore included as user studies in this thesis. Additional interviews are carried out during this thesis as well, but slightly modified to suit the thesis better.

2. Theoretical Background and State of the Art

2.1 Contribution of mobile devices in education

Research has previously been done on that shows benefits in m-learning (Sharples 2006, Kukulska-Hulme 2009, Sharples and Roschelle 2010, Ozuorcun and Tabak 2012, Trentin and Repetto 2013). These studies show common factors – mainly motivation and mobility – that are essential benefits for learning on mobile devices, but also mention restrictions and limitations to the devices. Mobile devices have the properties of being spontaneous, informal, personalized and ubiquitous (Miangah and Nezarat 2012). These properties can be seen as beneficial when it comes to learning. Research shows that these properties affect behavioral motivation in users, and motivation is a key factor to success in learning (Jovanovic and Matejevic 2014).

The most important benefit mobile devices have for learning is motivation. Among factors that determine effectiveness in learning is the importance that the user maintains a motivation to continuously learn. On mobile devices, motivation is mostly intrinsic, as users initiate their own learning (Trentin and Repetto 2013). Intrinsic motivation is more effective in a learning context than extrinsic motivation. People enjoys the activity, are more creative, and information is better perceived by the receiver (Ryan, Deci et al. 2000). The target group in this thesis may vary on the intrinsic and extrinsic motivation, as language learning is a requirement for them. Some may have the will to learn on because they want to, but others may see it as compulsory work. Mobile devices benefit from intrinsically motivating the user by having the ability to be accessible at all times, which promotes the opportunity of voluntary initiative to perform tasks. In order to achieve intrinsic motivation, three needs will have to be fulfilled – competence, autonomy and relatedness (Ryan, Deci et al. 2000). These needs are part of a concept called self-determination theory. Mobile devices fulfill these needs by granting rapid feedback and adaptive instruction

(competence), being learner-paced with tactile features (autonomy), and the presence of an audience and possibility of collaborative learning (relatedness) (Trentin and Repetto 2013).

A way to achieve additional motivation is to use gamification as a design method. As opposed to *serious gaming* that are games designed for learning, gamification uses only the *mechanics* of games and implements them to a product or service that is not a game. Mechanics of a game can be explained as the pieces that make the game what it is. By using a simple example of a game, the mechanics can easier be explained. Yahtzee, for example, contains the mechanics of randomness and points. Randomness is when the user throws the dices, which creates excitement. The points are collected based on the eyes of the dices. If a user has more points than others, the user may feel superior which creates a satisfactory feeling. The gamification approach in design is used to support user engagement, increasing user activity, social interaction, and/or quality and productivity of actions (Hamari, Koivisto et al. 2014).

Common m-learning may weaken the teacher and student relationship, which means the teachers will not be able to affect the student studying or his/hers motivation (Ozuorcun and Tabak 2012). Gamification of the application may however mend this disadvantage to a certain degree by offering the motivation for the user to come back, either by getting a higher score on a task or collecting more badges and points on their user account. Another problem with gamification is mentioned by Hanus and Fox (2015), as they performed a longitudinal study comparing a group of students attending a gamified course versus a similar group attending a non-gamified course. The students who were attending the gamified course were reportedly decreasing in intrinsic motivation and satisfaction over time. What they did fail to mention was the reason for choosing the specified gamification mechanics they used (badges and leaderboard), except that it had been used in previous research. They admit that it is possible to incorporate mechanics with clear learning objectives in mind, as opposed to what they did. The gamification mechanics chosen will have to suit the context of the content (Zichermann and Cunningham 2011), or else it will be a pointless adding of mechanics just to add gamification. Hanus and Fox's study was also conducted in an analog environment with digital rewards, where gamification was constantly present, not as a supplementary arrangement that the user could access voluntary.

The most used language learning applications are using gamification as a way to motivate users to continually keep interest and motivation in their applications. The effects of gamification are reported to affect motivation rather than attitude or behavior (Hamari and Koivisto 2013). A study done on the language learning application “Duolingo”, states that 34 hours of use equals one university semester, and that motivation was the main factor for higher effectiveness (Vesselinov and Grego 2012). This study was only conducted on Americans learning Spanish. Both the American and Spanish language belong to the Indo-European language family, which means they share similarities as the Latin alphabet. The Hispanic population in the US is 48.4 million, which makes it the biggest minority or ethnic group in the nation (United States Census Bureau 2010). With the amount of Spanish speakers in the US, it will be reasonable to suggest that some of the participants in the study already were familiar with some of the language. It is therefore expected that the results would differ from other nationalities learning Norwegian. The study also showed indications to that the application was more effective on a lower level of knowledge of Spanish, rather than a higher level, and its effectiveness decreased exponentially with the knowledge of users. That is however normal when learning a foreign language, as a more advanced level of the language is more demanding to learn. The basics are usually words and common phrases, while a more advanced understanding requires the learner to e.g. edify sentences correctly, conjugate verbs and develop their vocabulary (Modern Languages Division 2001).

2.2 Interface Design on Mobile Devices

There are studies on how to design for m-learning applications (Seong 2006, Ozuorcun and Tabak 2012, Segaran, Ali et al. 2014). These articles mention different guidelines and requirements for the term m-learning. The limitations of mobile devices need to be taken in consideration when designing a user interface. Seong (2006) mentions restrictions regarding screen size, input availability, battery life, processing power, resolution, memory, and bandwidth. These technological features have however been improved exponentially over the last few years, except input availability and physical screen size, as these components are essential to limit for the device to be categorized as mobile. He also mentions that mobile phones

decreased in size over the years along with screen sizes. Contradictory to this evolution, most of the recent phone companies have removed the physical keypad from the device, and started using touch screens (Sharp, Rogers et al. 2015). This affords all interaction to happen on the screen itself, which frees space for a larger screen. Tablet computers are also a product that has gained a place in the market. These devices allows for an even bigger screen, and offers the possibility of more organized information and computing power. Both tablet computers and mobile phones are regarded devices under the term m-learning.

The concept of m-learning addresses that learning does not have to be in an educational context, as in a classroom or through a physical interaction with another teacher or student. Pereira and Rodrigues (2014) mentions state of the art technology that use SMS and MMS on mobile phones as a form of feedback through applications. As this is an interaction with another human being, it is good in terms of feedback quality. It was however used as an alternative, mainly because there is a limited wireless society where the study was conducted (Rural India). Contact through the application – as a messaging system or chat – would make the application more suitable for both kind of mobile devices. Mobile learning means that a user can learn in any context, using an ubiquitous device – anywhere and at any time (Hashemi, Azizinezhad et al. 2011). An aspect that is important when using self-reliant solutions is technological adaptation. Direct teacher to student contact cannot be expected when being used as an out of context solution. Optimally, a learning system should detect what learners are struggling and succeeding with and adapt its tasks upon the user's current knowledge. Evolutionary learning environments are a proposal which arranges students in groups based on "type" (Pereira and Rodrigues 2014). In light of the target group, the type could be which knowledge level the student currently is attending at the learning center. However, when used in conjunction with a curriculum, a fully automated system should not be deciding quality control. With a mobile application, a backend with surveillance and control of the application's content can be accessible for the teachers. Like the current e-learning applications used in Norway – a system for the teachers to monitor detailed student statistics and generating of tasks. This way they can see progress of students, as well as appoint suitable tasks based on the level of knowledge. This should be essential, as the learning program in Norwegian learning centers divides the classes based on knowledge levels.

Design choices for a learning application should be done in accordance with reducing the cognitive load of the users, so that they process information more fluently (Seong 2006). The importance of minimal cognitive effort in educational applications is higher than in other applications, as the goal is to learn the content, and not how to utilize the application. Raymer (2011) supports the argument of minimizing cognitive effort, and proposes that a course within e-learning should be divided into modules and further into objectives. By completing individual modules, a course is ultimately completed. Structuring e-learning this way makes the users learn and practice new skills incrementally. What variety of tasks in the application will also need to be considered, as any additional mental effort may hinder learning (Chen and Huang 2014). This is further backed up by Segaran, Ali et al. (2014), but they argue for cost- and time-saving regarding downloading of content and launching of the application.

Much of the research that has been done is outdated in terms of dealing with devices that may not be as relevant due to the global market share today. An example is the *SymbianOS*, which at the current time is only at 2.78% in terms of global use. Google's *Android OS* and *Apple's iOS* is at the time primarily the biggest operating systems with an overall usage of 78,11% in the global market (Figure 1). The research also focuses less on the user interfaces of the applications, and more on the devices themselves – what limitations and possibilities they have for learning. The suggested solution should be designed with both Android and iOS operating systems in mind, as these are the most used by a large margin.

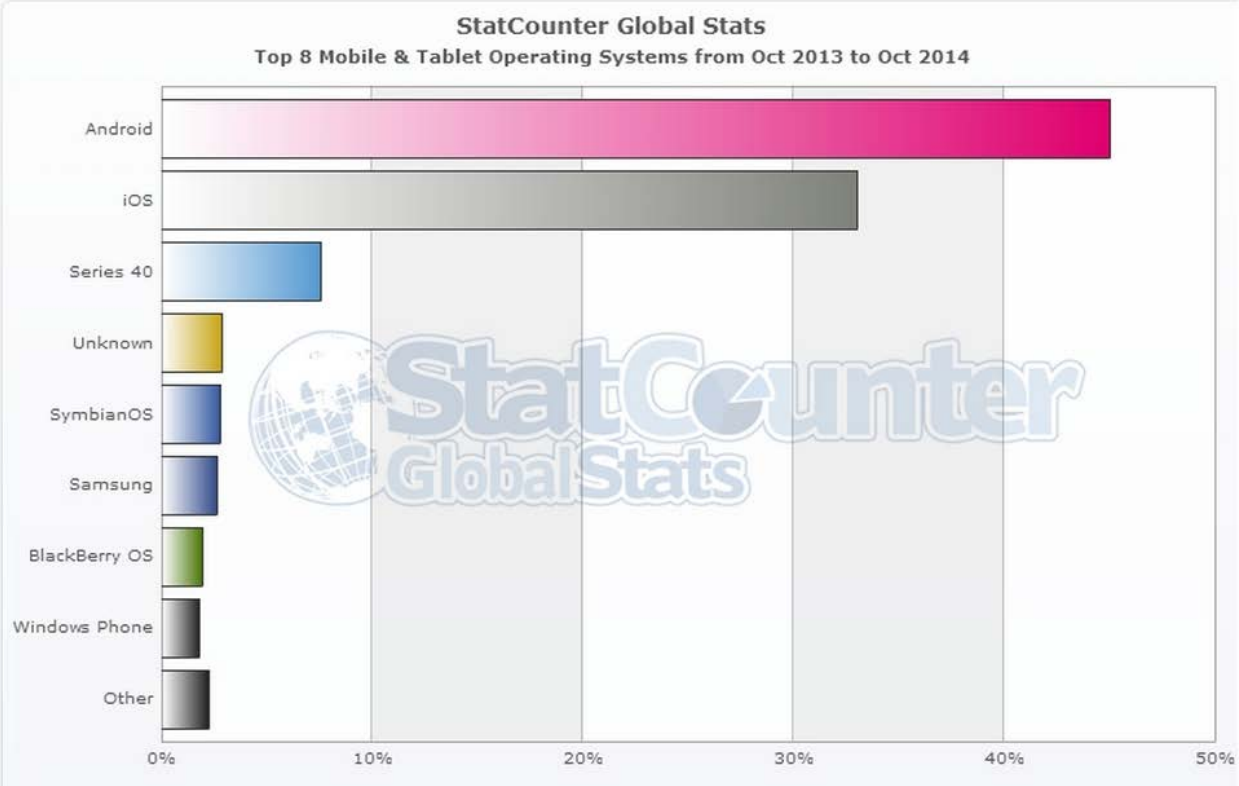


Figure 1 - Usage of different mobile operating systems on mobile devices (StatCounter 2014)[3]

2.3 Previous Use and Effects of Gamification

Gamification in learning already exists, although with divided results. Most studies have aimed at mapping learning outcomes and results. Little empirical work have previously been done on gamification and validation of its effects in non-game contexts. Authors state that further work is needed on effects of gamification on both context of use and demography of users (Li, Grossman et al. 2012, de-Marcos, Domínguez et al. 2014, Hamari, Koivisto et al. 2014, Seaborn and Fels 2015). This section will look at some of the previous studies done on gamification.

Hamari and Koivisto (2013) studied a workout application containing gamification mechanics. The application was heavily focused on the social aspects of gamification. The users perceived usefulness, enjoyment and playfulness tended to diminish over time. Achievements, levels and

[3] URL: http://gs.statcounter.com/#mobile+tablet-os-ww-monthly-201310-201410-bar

points were used as gamification mechanics in the application. The results indicated that only by age, ease of use decreased as a reported benefit. Hamari (2015) also experiments with badges used on *Sharetribe* - a website trading service. A longitudinal study was conducted where one year, the website was without badges and the next year badges were implemented. The study concludes with that users are generally more likely to interact and use the application more often when the badges are present. Since the application contains a social aspect, the result could be affected by how users want to maintain a high social status in the web community. Social networks – which are also a place where users strive to keep a good self-image – have been put up against gamification in another study. Social Network outperformed gamification in several courses. The conclusion is that one approach is not better than the other, but should be used where they may serve a purpose. There are some studies on gamification in learning as well, but with different contexts than what is intended in this thesis.

In e-learning, several studies have been conducted on gamification. McDaniel, Lindgren et al. (2012) used badges and achievements in online e-learning courses. Students reported mixed responses to the use of badges, since some of these badges were hard to get and affected the grades of the course outcome. In general, the study proclaims that the designing of badges could be done differently, and hidden versus visible badges could be a factor for the results. Another study on badges and achievements was done by Denny (2013). Student results were not improved from including badges. Student activity was however higher for when badges were implemented. The study concludes with that badges had a positive motivational effect on the participants of the study. Domínguez, Saenz-de-Navarrete et al. (2013) used gamification mechanics as levels, challenges, badges and leaderboards in an e-learning application to increase motivation of participants. The study concludes with that it seems it has potential to achieve the motivational effect, but the response from participants was mixed. The results of the participants were overall better, but the performance and participation was poor. There was also problems with the ease of use of the application. It was likely the factor for the poor performance and participation.

Attali and Arieli-Attali (2015) studies the effects of points in a mathematics application. The effects that are studied are accuracy and speed of responses in a 3D game. The results show that there were no more accuracy in responses, but in speed there was an improvement. Points and

badges can be seen upon as extrinsic rewards, as was mentioned earlier. They do in other words serve as additional motivators for learning. Su and Cheng (2013) studied the use of a mobile game-based learning system with achievements. The students in the experimental group were satisfied with the learner-controlled pace of the learning process, and they thought the gaming mechanics in the structure of game play were organized and useful for assisting them to learn. Perry (2015) studied a Language Learning gamified system on mobile using GPS, augmented reality and with quests as a gamification mechanic. This was used to discover the campus of the university while learning French. The preliminary results were showing that the application had potential and were motivating its users.

To design for a game, primarily the players or users intended to play the game, will have to be considered. A widely referenced author named Richard Bartle has defined different player types of games. These player types are named *Achievers*, *Explorers*, *Socializers* and *Killers* (Bartle 1996, Zichermann and Cunningham 2011). The player types can be shortly summarized:

- Achievers – Players who like to achieve rewards through something in games and game-like contexts.
- Explorers – Explorers are players who like to explore every aspect of a game. They like to discover things and value the experience of a system.
- Socializers – Players who like to socialize with other players.
- Killers – This player type likes to win, but also display their skill. Winning in itself is not enough, which means something has to lose or be affected in a negative way.

Players are usually a mix of player types, and a player can inhabit several characteristics. Although, considering the goal of users learning a language through an application, the language learning itself has to be seen as the intrinsic motivation for usage. Since the user group are learners, achievers will mainly be the focus when making a prototype. Users would like to learn something, and by confirming their skill and progress through rewards is seemingly the best use in the context. Rewards also serve as an extrinsic motivation in this case. However, it is also wise to consider explorers as well, since users should be motivated to continue using the application. There may be an extension of motivation if these player types are in mind when making the prototype. Socializers are not a priority in this case. The application is supposed to make the user self-efficient, and not in need of any help. The social aspect of interacting with both teachers and

students is achieved when the students are at school. There are also limitations in the project that makes the social part hard to test through a prototype, since the prototyping tool does not allow it. Lastly, killers are not involved since they do not fit the context of the application. It is also the least populated group of the player types in general and will therefore not be considered at all. This is further backed up by Heeter, Magerko et.al. (2009) that justifies the statement by comparing education with the player types:

"*..So let's see how each of our player types measure up to this new purpose.*
An Achiever will do whatever it takes to complete the course. An Explorer will explore all that the game has to offer thereby covering the whole course. The Socializer will work with all the other players of the game but may not complete the course. The course will have nothing that will motivate the Killer to complete it. Achievers and Explorers are the only types of players valid for educational games." (Heeter, Magerko et al. 2009).

The conclusion after studying the results of these articles are that effects of gamification are greatly dependent on the users and the context it is used in, as stated by de-Marcos, Domínguez et al. (2014). Gamification is not simply a combination of game elements such as points, badges and leaderboards, but must also be viewed as an experience for the player (Werbach and Hunter 2012). An application for learning a skill will greatly stick out from an application for tracking physical exercises. That means different gamification mechanics should be used in different contexts, and that is what diversifies the results of current research. Gamification in learning should be treated differently than other gamification approaches. Learning in this study happens in a different context, and with a different demographic than the studies mentioned above. It will be a different starting point in the study, and may also produce unique results because of that.

2.4 Current state of the art - Application Review

This section will cover some of the most popular mobile applications for language learning that in the time of writing that is available on mobile devices. The applications were chosen based on top results when searching for "language learning" on Google's *Play Store* - the application market for the Android operating system (Google 2015). This section writes a general review of

the applications – how they are designed, and which gamification mechanics they use. Note that these applications are reviewed by using the versions available February 2015. They are in other words an updated version from when the previous review was done in the pre-study, which was during autumn 2014 in a course named “IMT4882 Specialization Course 2” at Gjøvik Unitversity College (Myhre 2014). One of the computer-based e-learning applications used at the learning centers is also reviewed in this section. A temporary license for access is acquired for the purpose of this thesis. The section will end with a conclusion of the reviews.

2.4.1 - Duolingo

The graphical design of this user interface is easily readable. Elements (icons and text) are separated with a good distance, and the colors are strict and limited within a color scheme. It is mainly blue and white that is mainly used throughout the application. The exceptions are feedback elements, as when the user answers correct or wrong on a question they get a green and red color respectively, which users most likely will relate to due to color-coding.

The learning consists of different categories that can be seen on Figure 2. These categories are part of sections that needs to be completed to unlock further sections. The categories are again divided into lessons. The lessons address different words that will be learned within the category. The lessons are finally consisting of different exercises where the words are learned. This relates to what Raymer (2011) says about reducing cognitive load and incremental learning.

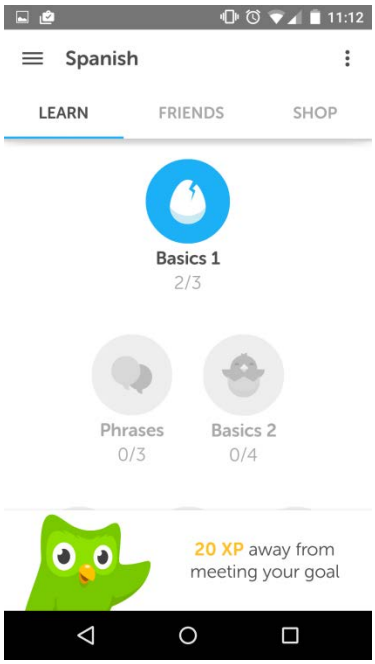


Figure 2 - The categories in the Duolingo application

The exercises are all with auditive feedback. The language that is being learned is pronounced as a word or a small sentence. There is also speech recognition where the user speaks a sentence or a word into the microphone, and the software recognizes the word. Other exercises are translation of words with or without alternative answers, listening to words and type/choose from alternative answers, and spelling of words. Different exercises can be seen on Figure 3.

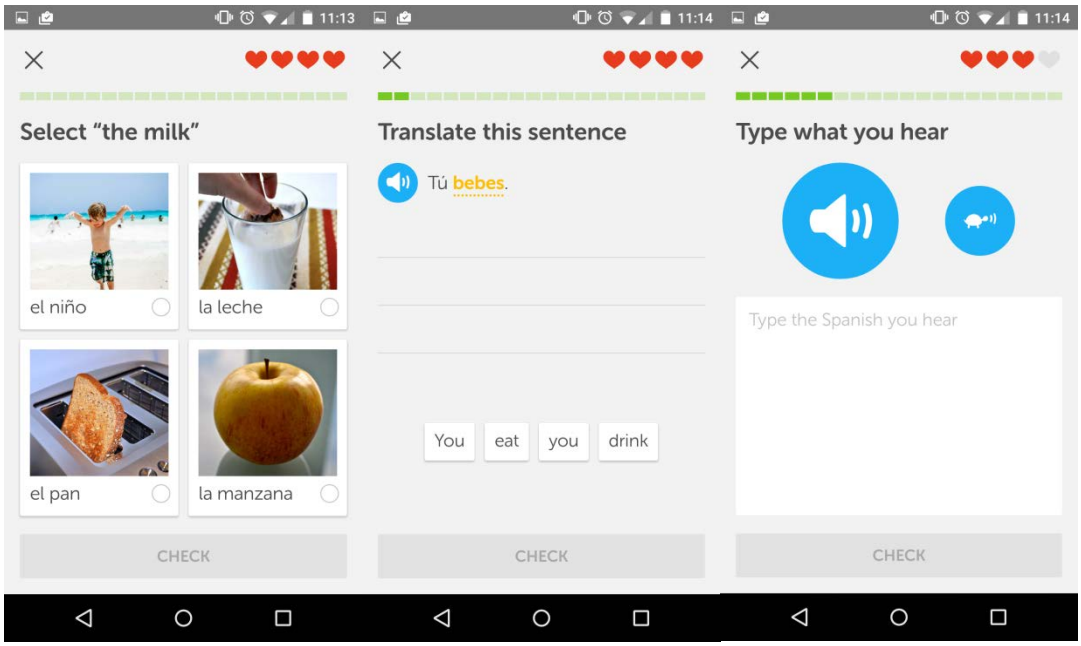


Figure 3 - Different exercises in the Duolingo application

The clear gamification is what makes this application stand out. During each lesson, the user has four hearts as seen on Figure 3. If a mistake is made, one heart is lost. When all hearts are lost, the user will have to restart the lesson. Success is also measured as XP, or experience points as the abbreviation is used for. The better performance by the user during a lesson, the more XP the user receives. After certain amounts of XP earned, the user will gain a level in the language being learnt (e.g. 25 XP rewards level 1, 50 XP rewards level 2 etc.). This is displayed as a badge to the user. The application suggests a daily XP goal (Figure 4), and is a motivation factor for the users.

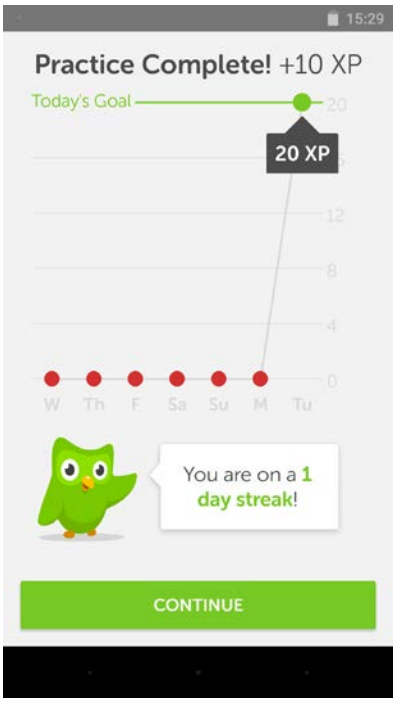


Figure 4 - Suggestion of daily XP goal

The weakest words of each category are remembered once a category is completed. That means the user can go back even though it is completed, and improve their skills within the given category. This is a great method for repetition of the language.

Gamification mechanics:

- Duolingo uses XP – points that are used for measurement.
- Player lives or attempts. This application allows the user to fail an exercise.
- The application contains levels, which shows progression and status of the user.
- Progression bar within exercises.
- The application contains a currency (Lingots) which can be spent on a shop.
- Unlockable rewards. The shop contains unlockable rewards by spending the games currency. Some rewards are purely aesthetic, while other rewards have a function (extra language lessons and refilling of hearts).
- The application uses the social aspect of adding friends that can follow your progress.

2.4.2 - Memrise

This application is rich in terms of content. The user can choose between extensive amounts of languages because of this. It does not however seem, as there is any sort of quality control of the content. That means the users cannot be sure about if the content is valid, since it is user-generated. The users have in other words collaborated to make the content of each of the courses.

The interface design is constructed more like a CMS (Content Management System). The creators of the application provide a template, while content creators add content. This content can seem different for various exercises, and the design can feel like it is inconsistent. This makes the interface feel a bit cluttered. Images can be blurry and stretched, which gives the application a sense of not being professional. This can further be a questionable aspect that gives the users a reason for considering other alternatives of language learning applications.

The feedback is similar to other applications. Red is signal for failure, while green is for success (Figure 5). The exercises are depending on that the user memorizes the words that are presented, and mix these words in the future exercises. The application does a good job in explaining the words that are presented by having the option of showing several images as help to describe the meaning of the word. As the content is user-generated, some words may lack images, and the quality of each course may vary. At the end of each course, the user gets results summarized with how many new words that is learned, how accurate the course is completed, as well as how many points that was collected.

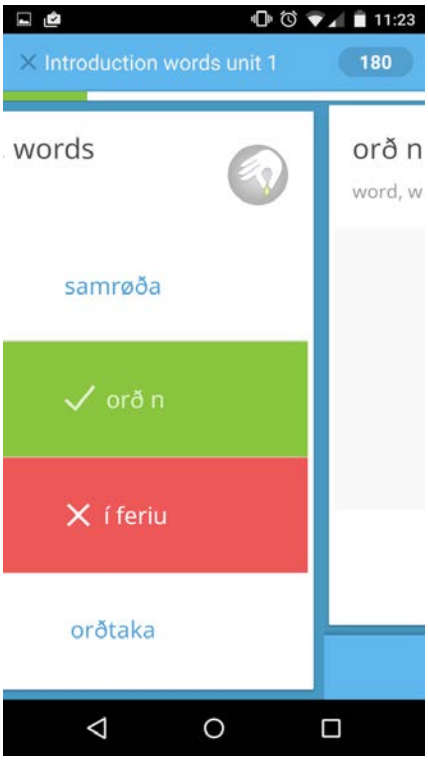


Figure 5 - Right and wrong in the Memrise application

There is also a review function which will activate notify the user after a given time. It collects words that users may have forgotten due to practicing it a long while ago.

One feature that makes Memrise stand out is the visual representation of progress in form of a plant that will grow once the user completes memorization tests (Figure 6). Users may find this feature as a personal challenge that will motivate them to continue completing tests, regardless if learning the language or growing the plant is their main goal.

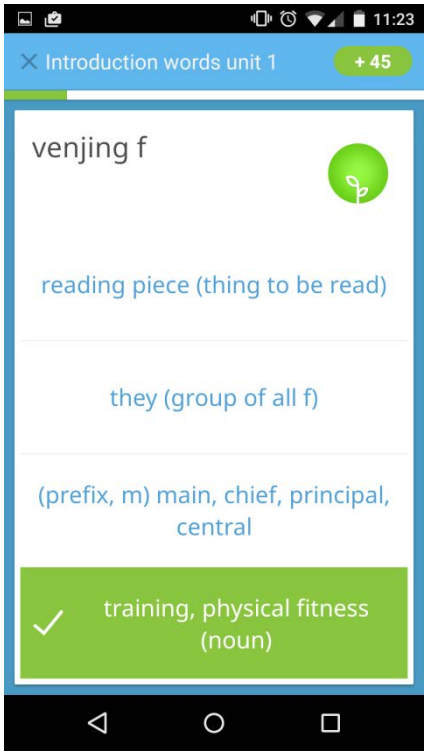


Figure 6 - Example of extrinsic motivation during language learning in Memrise: Growing a plant.

Gamification Mechanics

- Points depending on performance. There is also a daily goal much like in Duolingo.
- Highscore list among friends in the application.
- Grow the plant. The plant signifies your language skill, and is a symbol of progress and maintenance.
- Progress bar to show how far the user is through each exercise.

2.4.3 - Babbel

This application has a very simple design, and very limited options in terms of navigation. When a course first is started, the application remembers and continues where the user left off each time the application is opened (Figure 7). By only navigating with one interaction, the learning can start. The application does in other words inhabit great efficiency that the user can take

advantage of. The learning itself is in focus, and the interface is kept very clean. This makes the design intuitive with the user's goals in focus.

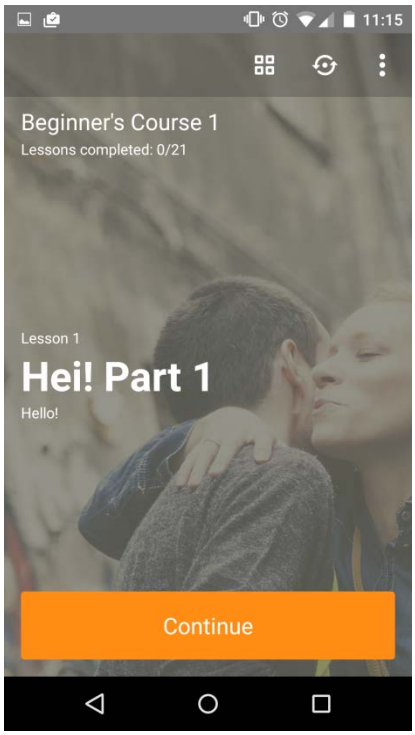


Figure 7 - Start-screen when course is already selected

Babbel allows the user to use the microphone in exercises to use speech recognition as a tool for learning. The creators of the application realises that the speech recognition still is not fully reliable, and lets the user skip it if desired. Other than that, the exercises is similar to other applications - Read and fill in missing words, speech recognition, choose the right translation(with images and audio as cues) and spelling of words with both the whole alphabet and the letters that are used as input.

The courses in Babbel feel more like an actual language course. Instead of only having exercises, small facts and information boxes about the language is displayed between and during exercises. This gives the application a feeling of being more fulfilling when wanting to learn a language. All exercises are read by a voice once completed. The auditive feedback is very similar like Duolingo.

Courses in this application have restrictions in terms of what is available to the user. As opposed to Duolingo, Babbel requires the user to pay a monthly fee to get access to all the content. The first lesson is always free of charge to let the user get a notion of what to expect. No skill level is required to unlock further exercises. This can be a dangerous kind of freedom for the user, as too many choices may disturb the flow of learning and make the users navigate between courses in confusion.

Babbel has an additional feature that is unique to this application - the Review Manager that can be seen on Figure 8. Review Manager lets the user practice words that they failed at during exercises. It is a great way for the user to get an overview of and practice words that they have problems with.

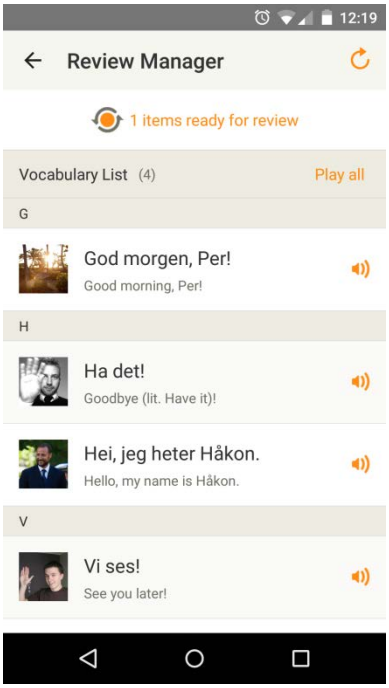


Figure 8 - Babble's Review Manager

The user gets a review of the exercise when it is completed that lists words the user needs more practice on. This is great for the user to learn incrementally.

Gamification mechanics:

- Progress bar of each exercise.
- Storyline in one kind of task where it is simulated a conversation. The user has to fill in responses in form of words for one of the characters.
- Points of completed exercise. (e.g. 46 of 47 correct)

2.4.4 - Fun Easy Learning Norwegian

This application is strictly for learning Norwegian. The user interface of this application is drastically different from the other applications. Instead of navigating through a hierarchical structure, all navigation happens on one screen by scrolling left and right (Figure 2.7). It is a bit unconventional. The effects of this approach are that the user will have to scroll through all categories to get a full overview of what the application has to offer. It also needs more interactions on the screen than the other applications when the user wants to choose a course and start learning.

The user chooses types of exercises. It can be seen on Figure 9 on the bottom line of the interface. The categories of learning are based on different themes, e.g. “Home” with subcategories like “Decorating” and “Toolbox”. The final bottom menu is for choosing the type of exercise. The exercises are shown on Figure 9, with the addition of “Choose Word” and “Listen & Choose”. When an exercise is completed, the completion bar beneath both the category and subcategory is slightly filled with a green color as an indication of progress. When all exercises are done, the bar is completely filled with green.

Figure 9 - Navigating the main menu of Fun Easy Learn Norwegian

The interface is heavily focused on visual cues during exercises. Most of the screen shows an image, while also playing audio of how to pronounce the word in question. This application has a thorough tutorial that explains how to use the application step by step (Figure 10). This is exclusive to this application among the others reviewed. It may be a sign of a user interface that is not very intuitive on its own.

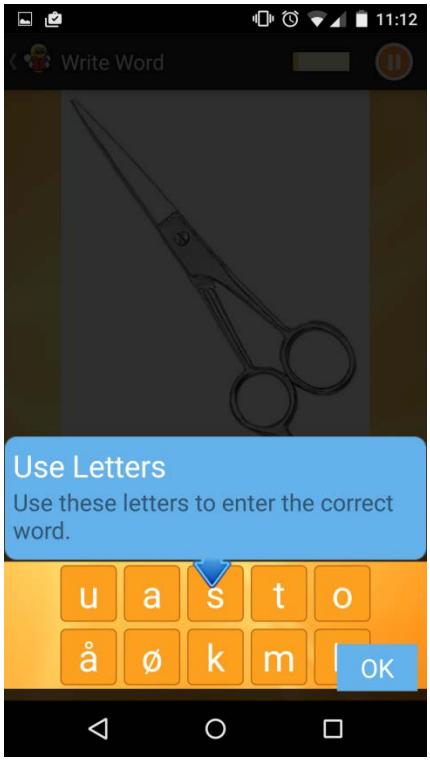


Figure 10 - Part of the tutorial of how to use Fun Easy Learn Norwegian

One additional feature that only Fun Easy Learning Norwegian has so, is the option to get help with exercises as seen on Figure 11. The user will get help with, e.g. a letter, in a spelling exercise. Since the application relies on users being self-sufficient, this feature can help them when they are stuck on something they find difficult.

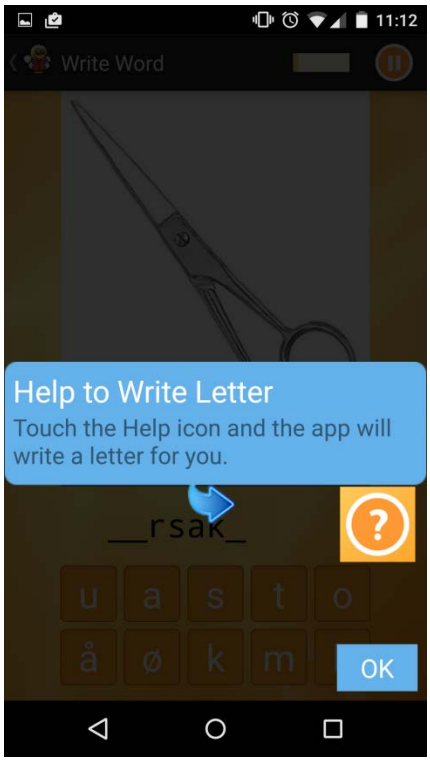


Figure 11 - The help feature in Fun Easy Learning Norwegian

Gamification mechanics:

- Points based on performance (flowers 1 to 3 per exercise).
- Collection of points (flowers)

2.4.5 - Migranorsk

Migranorsk is a web application that is developed by Norsk Interaktiv AS. It is developed for immigrants learning Norwegian – the same user group as this thesis focuses on. It was developed for the curriculum used in Norwegian learning centers, and is categorizing its content henceforth. Language levels that can be practiced through this program are ranging from A0 to B1, and can therefore apply to all students learning at both Gjøvik Læringssenter and Gran Voksenopplæring. The content is also grouped by themes, like “Living in Norway”, “Change of address

notification” and “Workplace ABC” that not only helps students learn the language, but also prepares them for daily life situations in Norway.

The exercises are widely different, and uses tools as multimedia and storytelling to engage the users. An example of an exercise is a user story of a person trying to get a job. This is presented through a video, and the student has to listen to what is being said. After the video is done, the student will have to answer questions to control that they understood what was said. An exercise that uses this example is seen in Figure 12.

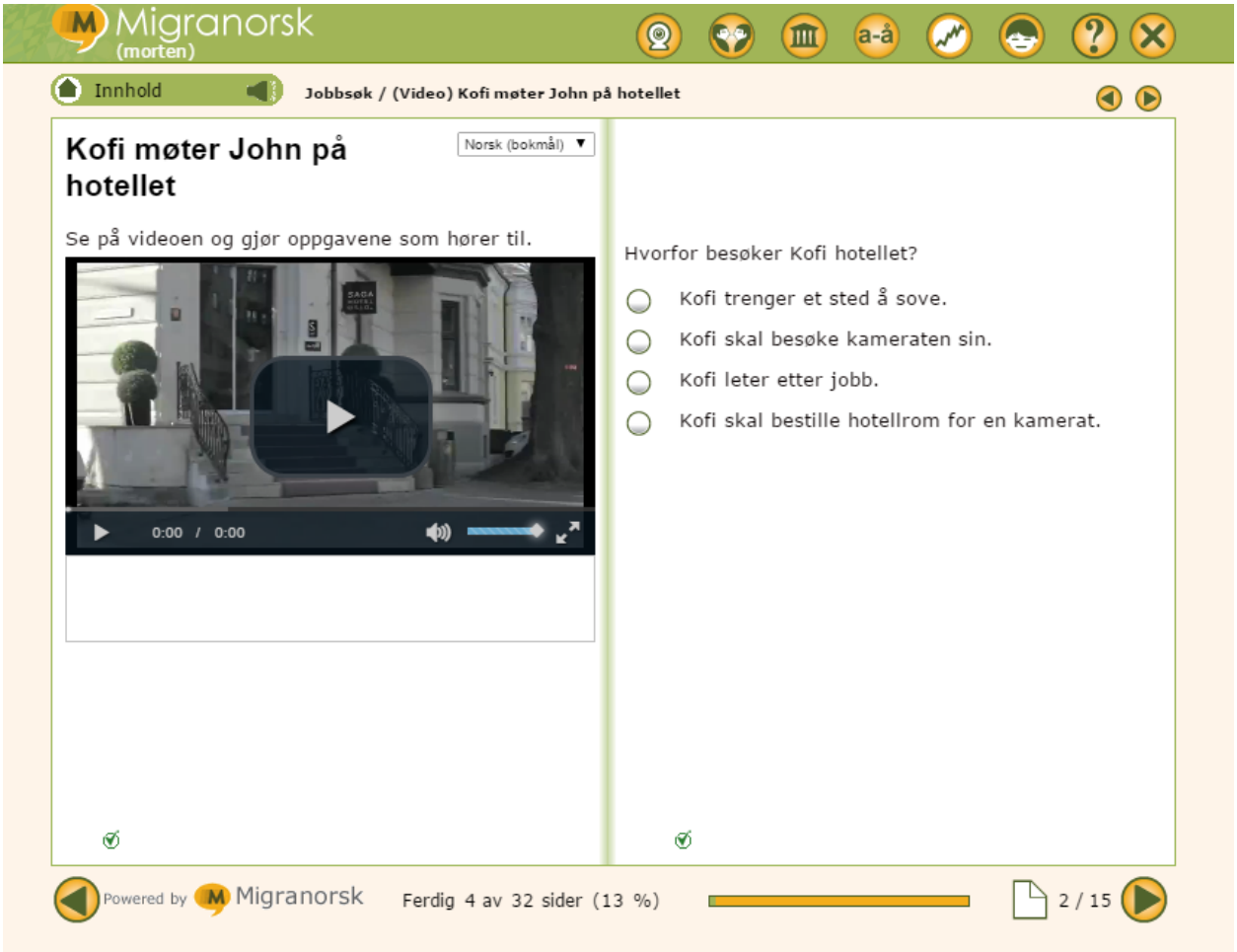


Figure 12 - Migranorsk exercise with video and storytelling

There are exercises that also use drag and drop-functionality and sound where parts of the page can be explained by a narrator. Other than these functions, there are mostly multi-choice answering and some free text answering that the user will have to send in for evaluation.

The statistics and results of each course that the user has attended can be seen in detail both by the user and an administrator (teacher). Progression, errors and time spent is measured in this part of the application. The student can improve on their own by seeing these results. An illustration of the statistics is seen on Figure 13.

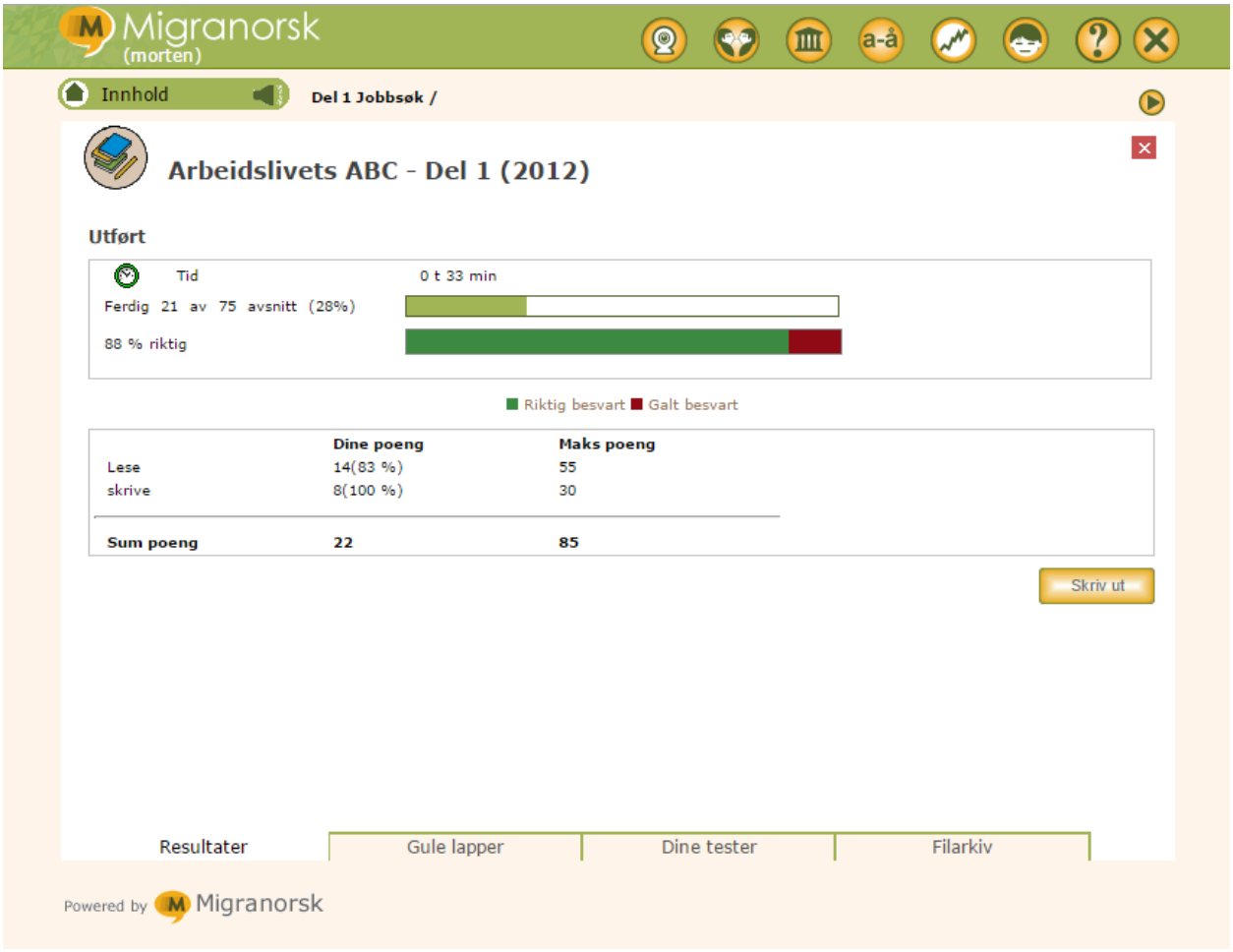


Figure 13 - Statistics in Migranorsk

When looking at some of the extra functionality this application offers, the glossary is considered useful. It is a simple glossary where the user can search for words and their meaning, and is easily accessible from anywhere in the application. The glossary can be seen on Figure 14. Migranorsk also offers social interaction with other students through the application interface. Both discussion and instant messaging is supported.

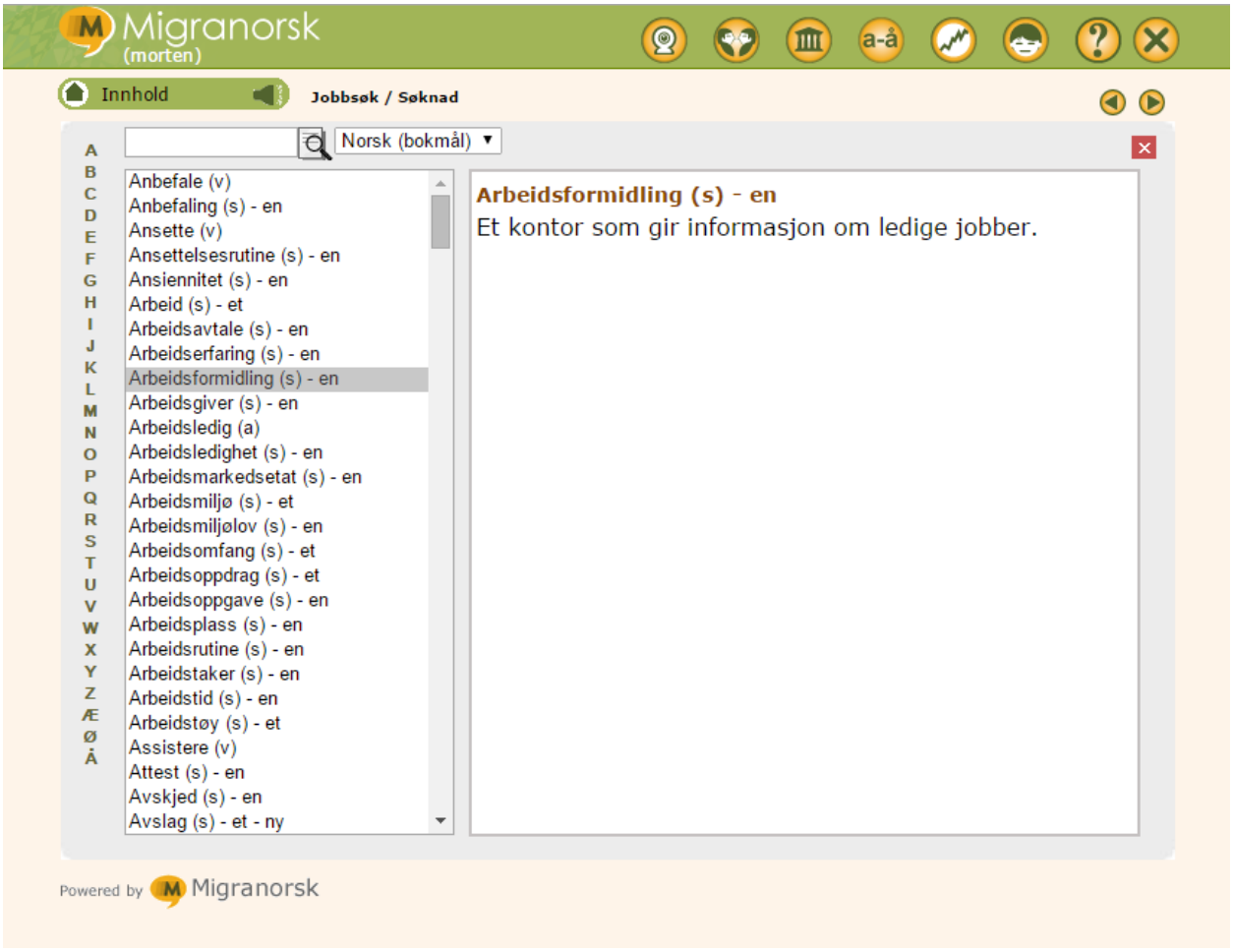


Figure 14 - Glossary in Migranorsk

Gamification mechanics:

- Storyline
- Time, points and percentage complete.

3. Methodology

This thesis will rely on research methods used in the area of Interaction Design to ultimately provide the data needed for this thesis. The thesis builds on previous work of the author, and uses interviews and questionnaires (Myhre 2014) from this work as a foundation for this project. These user studies were conducted at Gjøvik Læringssenter on both employees and students. This thesis continued by working with Gran Voksenopplæring to collect data at another learning center. The initial user studies consisted of conducting more interviews, and later led on to user testing of the actual users.

What was important for this project considering that there was an extensive amount of user studies and testing, is that participants were treated fairly and within legal boundaries. Following the guidelines of Data Protection Official for Research in Norway (NSD 2012) would assure that all participants of the project were treated with both ethical and legal consideration.

- When obtaining personal data from a participant, a written consent form will have to be signed by both parties – researcher and participant. The language of the consent form should be written considering such criteria as language, age, and cultural background. In case of collection of sensitive personal information, *Datatilsynet* should be informed of the project.
- The participant should to an extent be informed of the process and background of the research, as long as it does not affect the results of the study. This is normally a balance the researcher will have to determine, but would in this projects case not be a sensitive topic.
- The researcher will have to inform the participant about what information will be published, and where. The information should not be stored longer than necessary for the purpose of the project. If consent of the participant is collected, the information may be stored for further research. The participant will in case have to be informed of this.

- In case participants are bound to professional secrecy (in this thesis' case, teachers), the study shall be designed for the participants to express themselves without affecting their secrecy. They should also be informed to restrain from mentioning specific student's names or descriptive information (school score, personal information, etc.), as this is considered confidential.

All forms of recording and information that is gained from users or stakeholders are understood by the people involved, and preferably agreed upon in a written consent form. All participants will have to be informed what the information that is collected will be used for, and to whom the results will be viewed by.

The initial methods used were interviews. These methods were the *descriptive* and *relational* research methods in this thesis, while the experiment served as the *experimental* research method (Lazar, Feng et al. 2010). Lazar, Feng et al. (2010) describes three types of investigations in Human-Computer research, where two or more methods should be combined during research with human-computer interaction in focus:

● *Descriptive research*, which is investigations for gathering detailed data about the situation in the area of interest. This information is purely qualitative, and is usually gathered at the start of the process.

● *Relative research*, which is investigations that ties a cause and effect. This type of research can state that something relates to something else.

● *Experimental research*, which is investigations that can identify variables and identify causation of effects. This type of research usually involves control groups to be able to control the environment of unidentified variables.

Experimental research is in other words an investigation that had to be included in this thesis in order be able to identify causation of certain events and to bring reliable empirical data as contribution to the field. Rubin (2008) mentions that the methodology behind conducting a

usability test has originated from controlled experiments, and is therefore related to identifying cause-and-effect relationships.

The research methods was chosen based on what limitations that was set by the learning centers, as well as what data that needed to be collected during the lifecycle of the project. Several methods have been used, and this section will explain how and why the methods have been chosen and conducted. Figure 15 show how the methodology of this project and in what order it has been carried out.

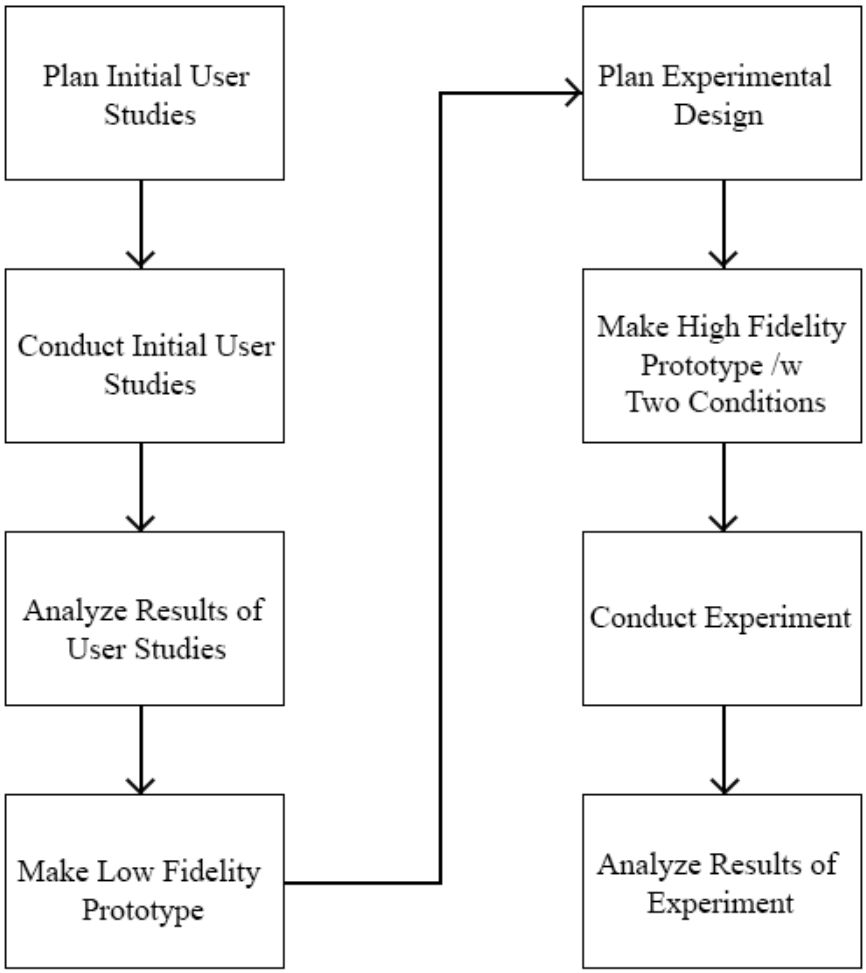


Figure 15 - A model of the methods carried out in this project

3. 1 Interviews

It is important to note that seven of the interviews as well as the questionnaire were conducted during a previous course (Myhre 2014), and not during this thesis. The interviews at Gran Voksenopplæring were conducted during the master thesis, and were used to compare to the results that were received at Gjøvik. The interview form can be seen in Appendix A.

3.1.1 - Goal

The earliest interviews were conducted at Gjøvik Læringssenter, and aimed for getting familiar with the current situation in adult language learning. The type of research that was first conducted, is called *descriptive research*, and is focused on collecting information on what is happening in the context (Lazar, Feng et al. 2010). The reason for conducting this interview was with several goals in mind:

- Get an insight in which procedures and methods that was used to teach students the language.
- Understand the demographic of the applications target group.
- Get a notion about how the daily routines of employees were, and potentially how a mobile language learning application could fit the learning plan.
- Get information about the tools used in e-learning, and how pleased both teachers and students were with these.
- What other tools that was used and which of them that benefits language learning the most.
- How skilled the average student was in using technology like PCs and mobile phones.
- Get a notion of what features that can be used in a mobile prototype.

3.1.2 - Target Group

For the interviews, the target group was teachers that taught immigrants the Norwegian language. This target group was preferred over the actual students because they could answer more advanced questions and provide richer qualitative data of how the actual teaching was done, daily routines, as well as the thoughts behind the current educational program. The teachers also had experiences from more than one class of students, and they knew more about the average student than the students themselves. There were in total eight participants with the title of teacher, two who were working as e-learning administrators, and one advisor. Three were men, and the rest women. Except for one ethnic Russian participant, all were Norwegian. They were all assisting students in learning the Norwegian language. Years spent by each participant in the profession averaged at 8.7 years, so most participants had broad experience with their work. Their average age was 42 years old, and most participants described themselves with a low competence in using digital devices.

3.1.3 - Recruitment

Recruitment was done by contacting teachers at both learning centers, which worked as intermediaries, and further made internal agreements with other teachers about being participants in the project. When asking the teachers to participate and help recruiting, a brief explanation of the project was done and the wording was laid out so that they understood clearly that participation was voluntary. It was also clearly stated that if the interviews were interfering with the lecturing, they would have to decline the inquiry.

3.1.4 - Setting

The interviews were conducted at Læringssenteret i Gjøvik and Gran Voksenopplæring. According to Saffer (2010), the best place for meeting when conducting a user study is in the natural environment of the subject. In that way, the meeting will be less artificial. The teachers were brought into an empty classroom where we could perform the interview without any disturbance. Two cases were an exception as they were conducted through a telephone call.

3.1.5 - Procedure

The interviews were performed as semi-structured. The term *semi-structured* means that a structured interview form will be followed in the interviewing process, but the possibility of follow-up questions can be asked when the interviewer wants (Lazar, Feng et al. 2010). This is a method that benefits when wanting to extract detailed data about a subject/area, and the researcher have pre-defined goals with what kind of information he/she wants to extract.
The interview objects were informed before the project that they would remain anonymous in the report and that the interview would be recorded, but only used for gathering detailed information post-interview. They were also informed that the results would likely be published when the thesis was done.

3.2 Prototyping

Prototypes are used for testing a product or service. Prototypes can be either low fidelity or high fidelity. Low fidelity prototypes can be made by using paper and post-it notes and simulating functionality. These variations of prototypes are usually not high in detail, and provide a cheap and fast alternative for testing. They are best used while still early in the design process to uncover the most general flaws (Sharp, Rogers et al. 2015, p. 392). While low fidelity prototypes usually are made for testing, the purpose of making them in this project was to get an impression of the graphical user interface as well as having a foundation for designing the high fidelity prototype. A sketch of the scenario can be seen in Figure 16. This project did not use low fidelity prototypes for testing for several reasons:

- To simulate gamification mechanics, a dynamic prototype will have to be used in order to show the points and badges judged by the outcome of each participant's performance.
- Low fidelity prototypes, while fast to perform, should also be tested on the real target group. This means agreements, planning and execution will have to be done. With the time frame of the project, valuable time may have been risked for little rewarding results.

- The tool that was used for producing the high fidelity prototype made it possible to make the prototypes in a relative short time period.

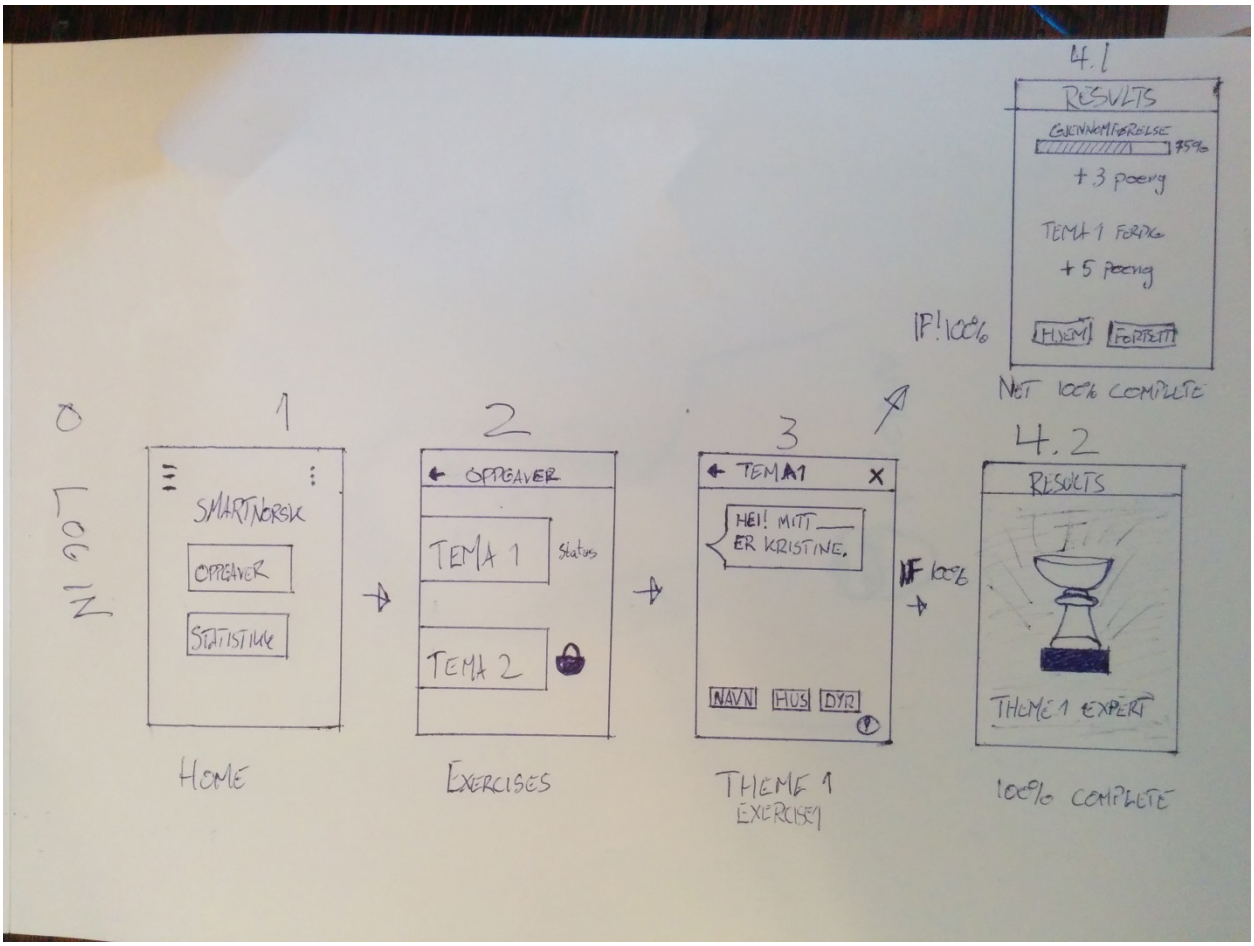


Figure 16 - Low fidelity prototype of a scenario made by sketching

High-fidelity prototypes are usually made with a high effort, cost and detail, and are made towards the end of the design process. Figure 17 shows the start screen of the final prototype, while both scenarios can be seen in Appendix B. The aim was to test with an appearance as similar to the proposed service/product as possible that further would render a more reliable user experience true to the final product (Sharp, Rogers et al. 2015, p. 395).

Figure 17 - Screenshots of the high fidelity prototypes made for testing

3.2.1 Goal

The goal of making a prototype was to test the variations of the application on users. The aim was to render the gamification mechanics as close to a real application as possible. With a prototype the hypothesis could be tested by making the users complete both conditions in an experiment.

3.2.2 Design Procedure

To make the prototype, the *Justinmind*[4] prototype tool was used (Figure 18). The software had a user interface that could be compared with Adobe programs, and was therefore relatively easy to learn and familiarize with. The tool also supported testing on mobile devices, which made it even more suiting for this project. By using the low fidelity sketched prototype, the design was both fast and easy to reproduce.

[4] URL: http://www.justinmind.com

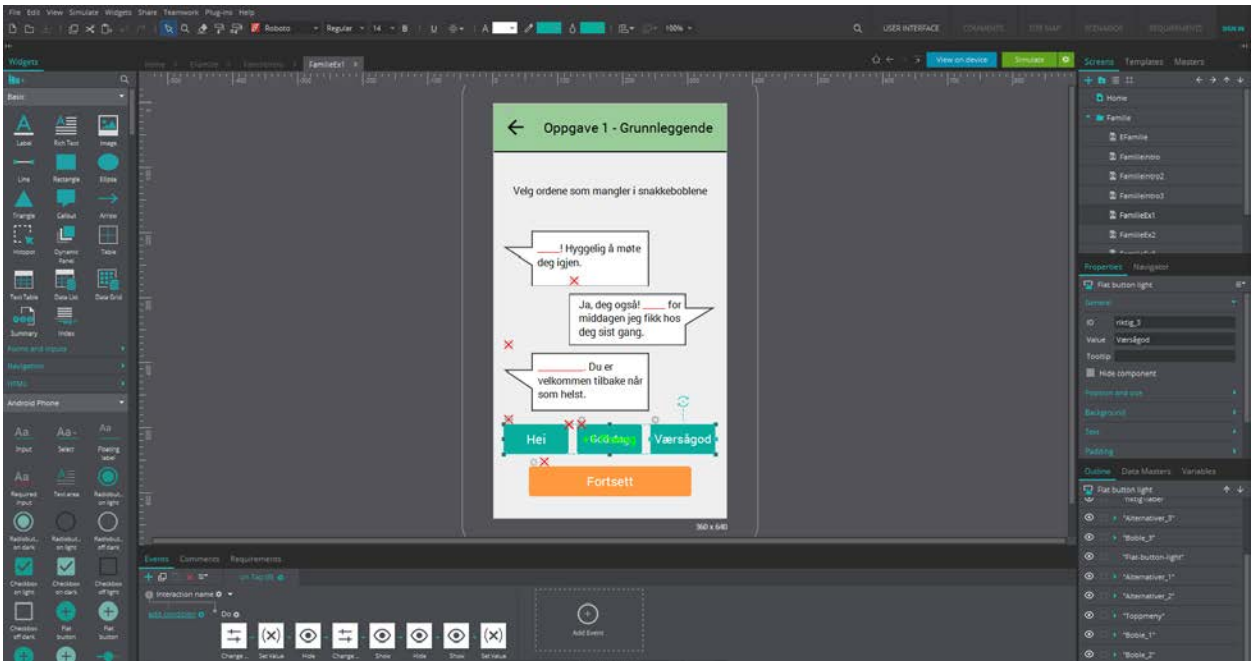


Figure 18 - The Justinmind prototype software

3.2.3 Content

The exercises were made based on exercises in e-learning (Migranorsk) currently used by some teachers in the learning centers. The reasoning behind basing the exercises on e-learning rather than m-learning was that state of the art applications use two languages in their applications; the language native to the user (the mobile device's system language) and the language that is being learned by the user. Both Gjøvik Læringssenter and Gran Voksenopplæring operate with solely talking Norwegian to the students and vice versa. To follow the trend of the learning centers as well as the e-learning software being used there, the mobile prototype would also exclusively contain Norwegian as a language. That presented some challenges, since most exercises in the state of the art applications were exercises based on translating between the user's language and the language being learnt.

The difficulty in terms of language skill of each exercise was made with the aim of participants being able to complete them with ease. With the help of a teacher, the exercises were verified to be easy for the language proficiency of the classes picked out for the test. The two conditions went through the exact same language exercises. The only variation in the two conditions was

the gamification mechanics, to ensure that the exercises would not interfere with the satisfaction of each user.

3.3 The Experimental Design

An experiment was performed to prove the hypothesis right or wrong. A classical experimental design needs according to Rubin (2008, p. 23-24), the following requirements to be accepted:

- **A hypothesis**. The hypothesis is stated at the end of this chapter.
- **Choosing a random sample group**. The participants were chosen directly by the teachers, and can be considered chosen randomly within the bounds of which test subjects that was desired.
- **Isolating and controlling variables**. As far as possible, controlling variables was done by keeping the exercises in the prototype similar and asking questions at the end of the test period.
- **A control group**. There was no control groups in this project, since all participants performed both conditions.
- **A sufficient number of participants.** By not using a control group, but rather use a within-group design on the experiment, less participants were required.

Fulfilling all these requirements except having control groups, the experiment was considered sufficient to carry out for the goal of this thesis. The number of participants was not low, but also not high enough to make the results statistically significant by only assigning one condition to each participant. Therefor by assigning both conditions to each participant, but in random order, the experiment performed were a within-group design (Lazar, Feng et al. 2010, p.42). The advantage of this design is that fewer participants are needed. Since two conditions are tested, half the participants are needed as opposed to a between-subjects design. Both learning effect and impact of fatigue are limitations of such an experimental design. The disadvantages mentioned by Lazar, Feng et al. (2010, p.49) were however considered of no consequence in this

project. The experiment did not consider the learning effect as a constituent factor, since the content is not what was in focus. Fatigue was also not a problem, as the experiment was planned to last between 20 and 30 minutes per subject. A within-group design should work well for this kind of project and will be explained further in this section.

3.3.2 Target group

The target group of this test was the user group the application was intended for - immigrants learning Norwegian. The target group had to be able to understand Norwegian to a certain degree, and was therefore decided to be at least at level A2 in both oral and written language level. The amount of knowledge of the Norwegian language was not an important factor for the experiment results, but for the participants to both understand and answer the instructions and questions provided by the researcher.

3.3.3 Recruiting participants

Recruiting participants was done by contacting employees at both Gjøvik Læringssenter and Gran Voksenopplæring either by phone or by email. Only the most proficient classes were asked to participate as the test was conducted in Norwegian. The person that was contacted was informed that participants only would have to attend if they were willing to. Knowledge of the English language was not expected of the participants. Even though recruitment was done by communicating with employees, participation by each participant in the test was voluntary.

3.3.4 Structure

The test went through several steps that will be explained chronologically below. To ensure each test was conducted similarly and all participants were treated the same, a manuscript was read by the moderator through the process. The moderator was also present throughout the process, in case the participant needed clarification or explanation of how to proceed. In Appendix D, the manuscript used during the test is displayed. The testing process is roughly summarised below:

- The participant was welcomed and explained shortly of the test to come. The participant was given a consent form (Appendix C) that explained what the data would be used for as well as how the experiment would be performed.
- The participant was faced with one of two conditions followed by a SUS questionnaire (Appendix E) which is a questionnaire explained more in detail in section 4.4. Which condition the participants started with was determined by a random number generator.
- Step 2 was repeated until both conditions and questionnaires were completed.
- At the end of the test a set of questions were asked to be able to identify if other trends or patterns were important for the outcome of the SUS questionnaire. The questions were strictly related to the research questions
- At the end of the test, the student was thanked for helping bringing results to the project as well as asked to bring in the next participant.

3.3.5 Setting

The setting was identical as during the interviews - at the learning centers. It was not important for the participants to test in this environment, since the usage of such an application should be independent of environments. It was however done to save the involved parts of as much effort as possible, as well as making the participants more easily available.

The lab setup was simple and required little of both space and resources. The setup can be seen on Figure 19. The moderator was close to the participant, but not so close that the participants felt like they were observed with every interaction they performed.

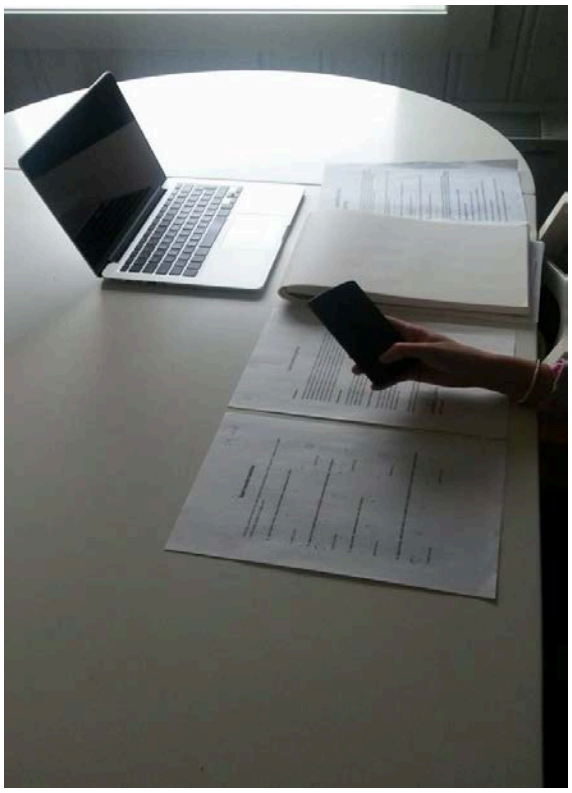


Figure 19 - Setting of lab experiment

3.4 Measuring Satisfaction - The System Usability Scale

According to ISO 9241, 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*" (Bevan 2006).

Usability goals are what Rogers et.al. (2015) defines as criteria that helps optimize the interaction users have with a given product or service. The term "usability goals" within interaction design is a word that recurs in different books and from different authors (Nielsen 1993, Rubin 2008, Lazar, Feng et al. 2010, Sharp, Rogers et al. 2015). They all refer to aspects of interaction that are possible to measure, but the goals themselves can slightly differ from author to author. What is consistent however, are the goals mentioned in the definition of

usability - effectiveness, efficiency and satisfaction. The only deviation is Nielsen's definition of "errors", which by definition is the same as effectiveness described by the other authors. The weight of credibility will therefore be in favor of Lazar et.al. (2010), since the authors are more focused on using these goals in a context of research, and calls these goals dependent variables. These goals will be noted and measured during user tests in case they affect the end result. Not all of these goals will be measured by quantitative numbers, but will be taken in consideration when performing the experimental test.

To measure satisfaction, a SUS (System Usability Scale) was used. As explained in the introduction of this thesis, satisfaction is a combination of *usefulness* and *learnability*. A SUS relate to these goals, as the questions provided addresses these areas. It is performed by answering a close-ended questionnaire. Figure 20 shows a standard SUS questionnaire (Brooke 1996).

System Usability Scale

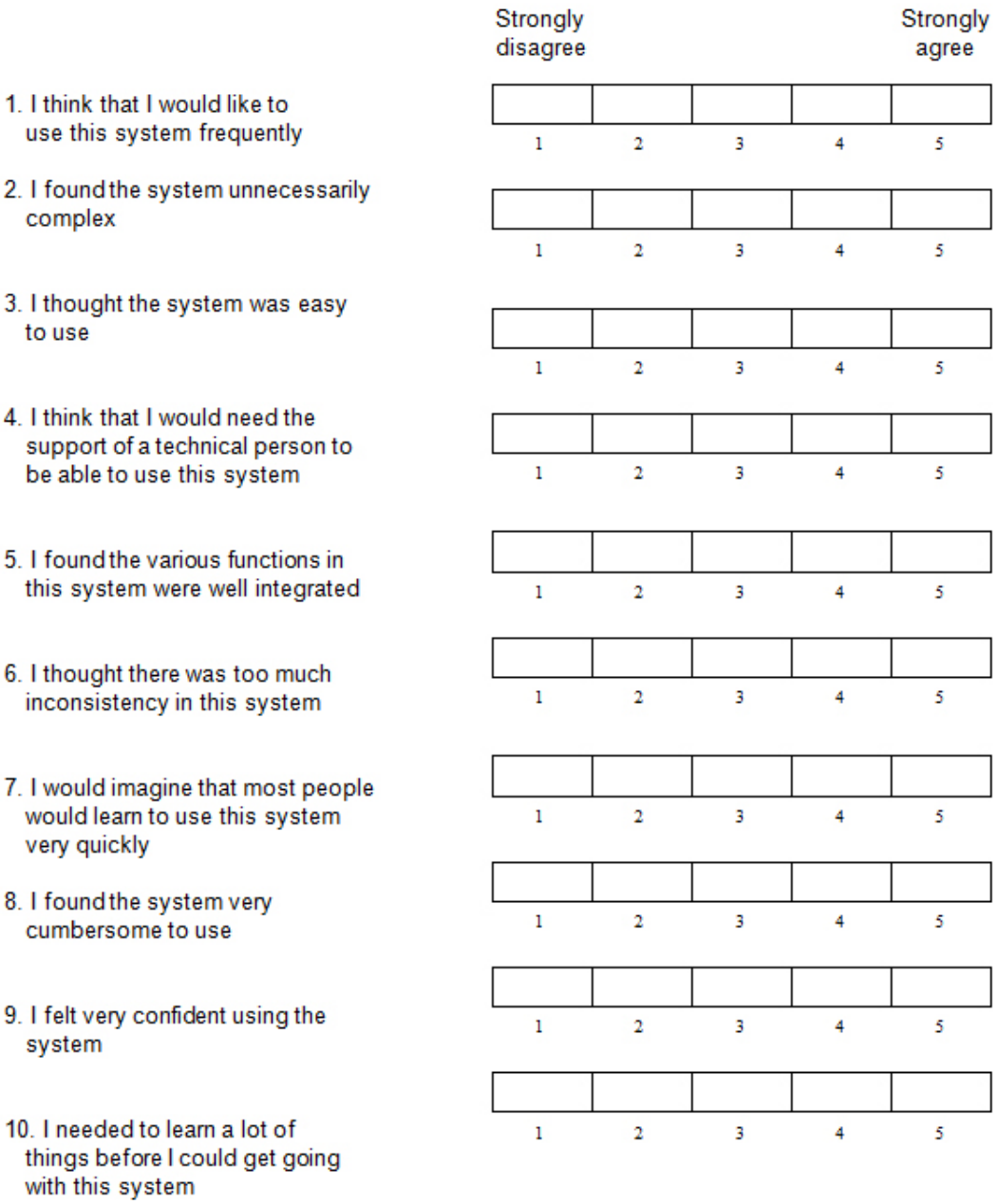


Figure 20 - A standard SUS questionnaire (Brooke 1996).

The SUS questionnaire is used for measuring satisfaction, and is a solely subjective rating determined by the participant. The questionnaire contains a total of 10 items where each item is rated by a Likert scale ranging from 1 to 5. Likert scales are numerical ratings, typically ranging from 1 to 5 or 1 to 7 (Lazar, Feng et al. 2010). However, to gather more precise empirical data from the users, the SUS (System Usability Scale) is proven to be both reliable and valid to gather data (Brooke 1996). The SUS was executed as a survey shortly after a usability test when the user still had the task in mind. The questionnaire is a part of ISO standard 9241 (Rubin 2008, p.194). SUS was chosen as a method that could produce quantitative and empirical data. To calculate the score of the questionnaire, the following formula was used:

1. Odd numbered item scores have to be subtracted by 1 (Score - 1).
2. Even numbered items scores have to be subtract from 5 (5 - Score).
3. All numbers have to be summarized and the final score multiplied by 2.5 (SUM * 2.5). (Jordan, Thomas et al. 1996)

The score can range from zero to 100 when the overall value is calculated. A higher score marks a higher satisfaction by the user. The questionnaire used in this project was translated to Norwegian, but contained the same definition for each item to keep the integrity of the questionnaire.

3.5 Hypothesis

The related research in chapter 2 shows that gamification mostly tends to increase the motivation and user activity of the players using applications with this approach. Since the results are mixed, gamification increasing user motivation cannot be presumed, but looks highly probable.
Therefore, the hypothesis of this thesis will be the following:

Hypothesis: Satisfaction of mobile language learning is better perceived by users as opposed to an application not using gamification.

4. Results

4.1 Interview

The interviews resulted in information regarding trends and usage of current e-learning at both Gjøvik Læringssenter and Gran Voksenopplæring. The interviews were recorded, and the most relevant information was noted. Relevant information was data that could be related to the research questions. That was also how findings were collected and analyzed. By grouping the answers by research questions, conclusions and frequency of mentions could be easier analyzed, and findings easier observed. The interviews were semi-structured, but follow-up questions were only used to ensure the questions were not misunderstood or that the interview object were not omitting any information. The follow-up questions used are defined as *probing* (Sharp, Rogers et al. 2015) so the structure of the interview was retained. That made it easier to compare results between each participant.

Data was analyzed by writing notes of relevance per participant during each interview, as well as listening through the recording to check if some information was overlooked. The research questions were changed after the interviews were conducted and an insight into the learning environment was made. Not all the answers to the interview questions will be directly related to a research question. The results of the interviews made a foundation for what to research, but have also been - along with the state of the art review - an important part in deciding the design of the prototype. The results provided in this section will be used for making the prototype that will be used in the experiment.

4.1.1 Interview Findings

The participants of the interview were asked about current usage of e-learning. The employees were using different software and different variations of these. It was both based on preference of

the teacher, and what language level the students were currently attending. The teachers also seemed reluctant to try software they did not have experience with, since they mostly reported they had a low technological competence. They had received some instructions on e-learning, but that was only for the currently most used application. Some software seemed to be covering a wide range of language levels. These applications were licensed and had been developed for such usage. It is important to mention that all of the applications here are web-based applications. None of the teachers or students that was contacted during the user studies had ever used a mobile application for learning Norwegian. The most used applications and their frequency of usage by participants can be seen on figure 21.

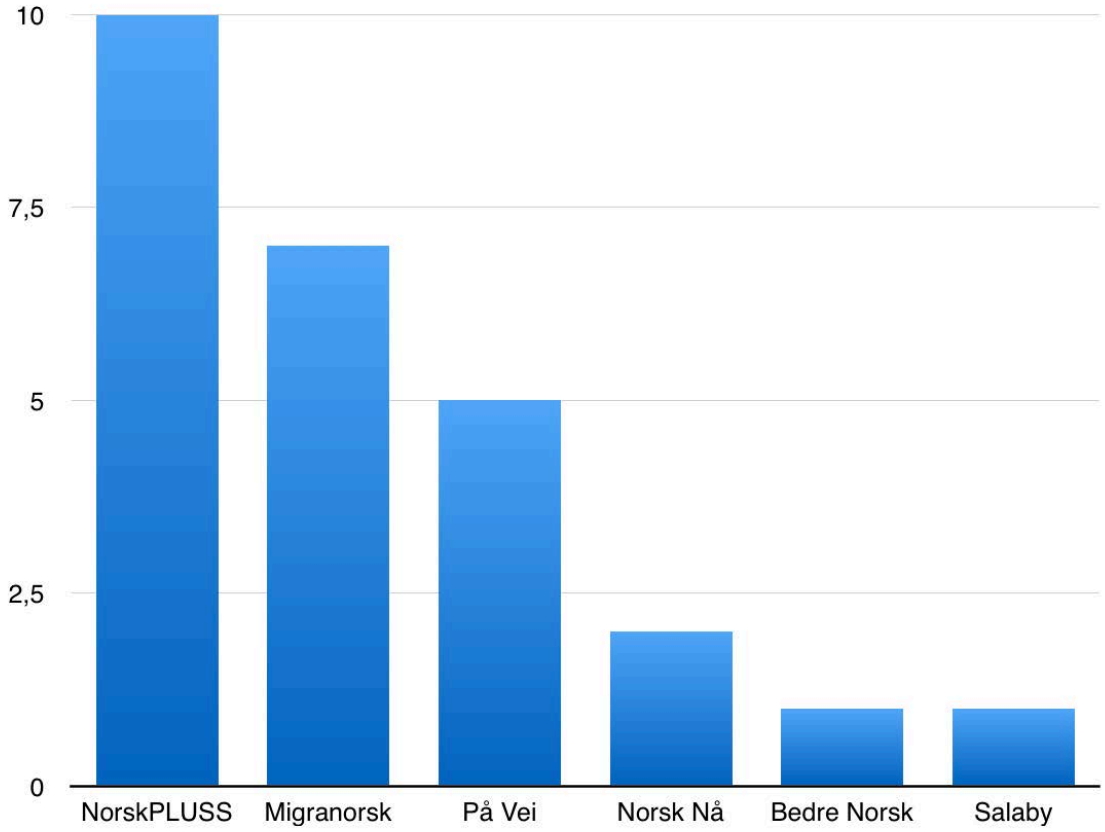


Figure 21 - Frequency of e-learning applications used by employees at learning centers

While several applications were mentioned, *NorskPLUSS* and *Migranorsk* were the two that were clearly favoured by the participants. As many as eight participants stated that NorskPluss

were their preferred application, while three stated the same about Migranorsk. Two participants also preferred both. In two other cases, other applications were mentioned, but the reasoning was that their classes were at such a low language level that the students would not be able to use either of the applications. Getting a license for NorskPLUSS was attempted during this project, but to no success. Access to the application Migranorsk was however acquired, and the application was reviewed in the state of the art section of this thesis.

When asked to explain what the benefits of the applications they used were, the answers were mostly that students could be independent. The question was formulated so that the participants could answer with what they thought e-learning contributed with as a supplement to traditional learning. Other benefits were mentioned as well, and the most frequent can be seen on figure 22.

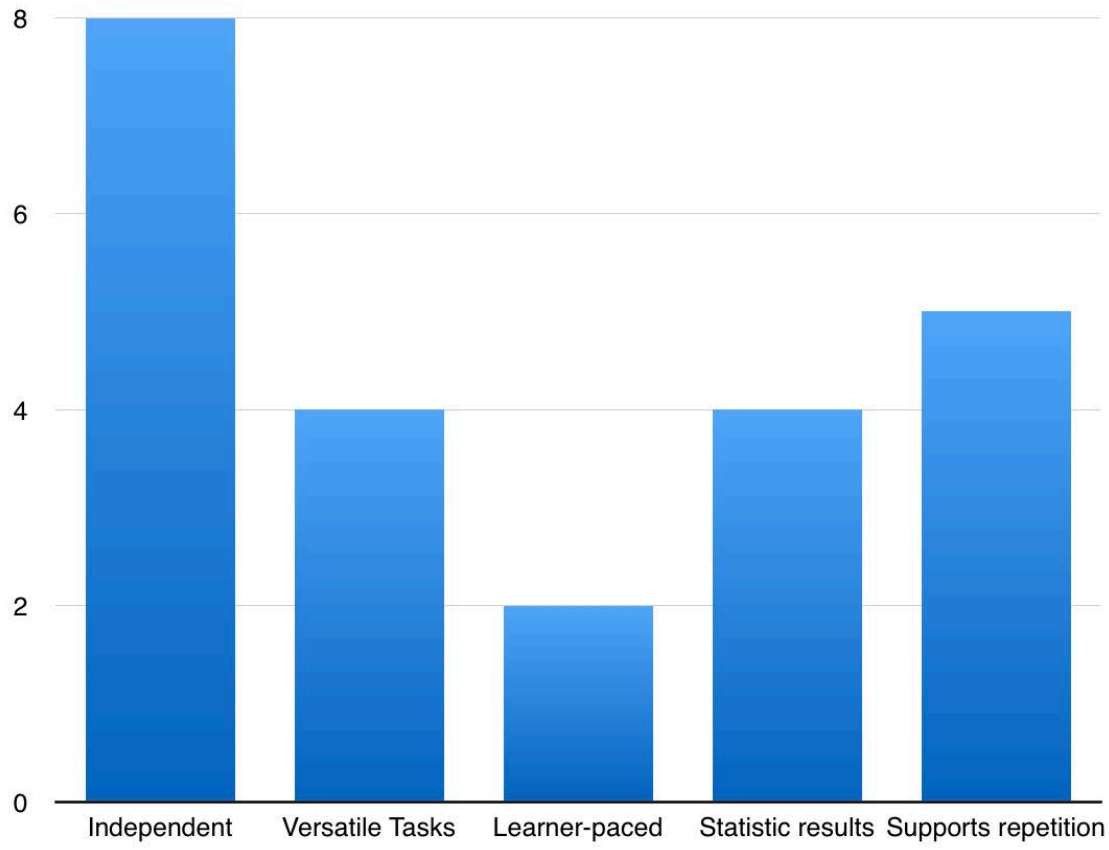


Figure 22 - Benefits of e-learning mentioned per participant in interviews

Other interesting findings were mentioned by several teachers as well:

- Six of 11 interview participants said there was a mixed competence of technology among students. Some were proficient and some were not. It was mostly a trend in their ethnicity.
- Eight of 11 said students are motivated during e-learning sessions. Three of the eight said the students are *mostly* motivated, which they explained with that the motivation is governed by their technological competence.
- Six of 11 mentioned that students currently used their mobile phone as a tool for help in learning. It is not allowed in class, but as it was only used for learning, teachers allowed it. The phones were used for Google translate as help in lectures, not with any learning applications.

Other than these findings, some interesting statements were expressed by some of the participants:

- "***Hey" I was not good enough at this task. I have to try it again***". A participant explaining what the students are thinking when explaining the benefit of repetitive tasks and motivation in e-learning.

- "***One of 20-26 hours a week is used for e-learning. This is because of a limited capacity and availability of equipment (Computers)***". A participant explaining that students get too little time for e-learning during school hours. This may affect how students approaches technology at home as well.

- "***I think something on mobile devices would be very helpful. Make sure users can grind and repeat tasks. 'Everyone' has a phone, and 'everyone' uses them. Phones are so crazy available. Anything can happen on a phone***". A participant talking excitedly about the potential of smartphones in Norwegian language learning.

4.1.2 Personas

To maintain an image of the primary users of the application, several personas will be made to more easily keep the user group in mind while designing the prototype. The personas will range from a typical user to edge cases (the most unusual users). The details relevant to the application are taken from interview and questionnaire results. The personas that are included are not real examples or real names taken from the user studies, but more of a mix and generalization of what was observed. These are the results from the user studies personified, and will be considered when designing the prototype.

Ianre Sakho

Gender: Female.

Nationality: Eritrean.

Age: 32.

Interests: Family, walking, arranges events with friends, singing.

Former education: 3 years of elementary school.

Language proficiency: Ianre knows some words in Norwegian. She can have a basic conversation with the teachers and some of the students.

Technological competence: Ianre's technological skills are very low. She owns both a smartphone and a laptop PC. Ianre knows how to send e-mails and use the Office-package to write and edit text files. In other words, it is the minimum requirements of what is obligatory to learn while attending adult education in Norway. On her phone, she only calls and sends SMS, which is the only way she knows how to use it. This is mainly to keep in contact with family and friends.

User story: Ianre is attending the Norwegian course on "Spor 2". Her class has just started using NorskPluss, and she is having trouble understanding how to navigate in the application. All the courses are available to choose from, and she is confused about which one to choose. She starts a course, and manages to do one third of it before they take a break. When she comes back, the progress is deleted, and she cannot remember which course she attended. Ianre is angry and irritated, and have to start all over again.

Adonis Bakas

Gender: Male.
Nationality: Greek.
Age: 28.
Interests: Photography, table tennis, nature, brewing beer.
Former education: Adonis is educated as a multimedia .journalist from Greece. Because he could not land a job in Greece, he was forced to look in other countries.
Language proficiency: Adonis is good at English, both written and oral. His Norwegian skills are not good when speaking or writing, but he understands much of what is being said and read.
Technological competence: Adonis uses technology on a daily basis. He is a hobby photograph, and is competent in the software Adobe Photoshop that he regularly uses as a tool for freelance work. He also often reads his e-mails and newspaper online. On his phone, he uses the camera and also likes playing games when he is on the bus or train.
User story: Adonis is attending “Spor 2”. He is usually doing e-learning at home, but as he has a busy daily schedule, he seldom has time sitting down and going through the recommended homework. He is using Migranorsk, which he thinks is an excellent solution for learning Norwegian. The only downside is that he has to power up his computer, access the webpage and log in – a process that makes Adonis less motivated to start learning during a busy schedule. “If only there was a faster way of access”, Adonis thinks to himself.

Da Chun Xiang

Gender: Female.

Nationality: Chinese.

Age: 24.

Interests: Video games, family, table tennis, web-surfing.

Former education: Da Chun has fundamental education from China, but is not specialised in any area of expertise.

Language proficiency: Da Chun knows some English, and is at about the same level in Norwegian. It is enough to keep a very basic conversation and understand simple sentences in both languages.

Technological competence: Da Chun has been using digital devices since she was a child. Her competence in using computers and mobile devices is coming natural to her, as it is part of her everyday life. She uses the internet on both mobile and computer in her spare time, and is very comfortable in doing so. She is also playing video games almost daily on both her computer and Playstation 4.

User story: Da Chun is attending "Spor 1". She has recently started using e-learning at school and as a tool for doing homework. She uses NorskPluss and attends the courses that have been suggested by her teacher. She is motivated for working at home but finds the design of this application kind of outdated, which reminds her of school and makes it boring to her. She wants an application with a fun design that can make her somewhat forget about school.

Images used in the personas were taken from Colourbox[5] using a license distributed from Gjøvik University College.

[5] URL: https://www.colourbox.com/

4.2 Prototype Design Choices

The prototype to be used in the experimental design was influenced by information collected by reviewing state of the art and information gathered through the qualitative user studies. The state of the art and interviews made it apparent that *structural gamification* (explained below) was the solution to this project. The gamification mechanics applied was determined to be points and badges. By not including more mechanics, the factors that may affect the outcome of the experiment were more limited and hence more controlled. As explained earlier in this thesis with Bartle's player types, the player types in this context are *achievers* and *explorers*. The prototype is developed and used by participants on an Android device, as Android is the currently most used operating system for mobile devices.

4.2.1 Research Question 1: Gamification Mechanics

All of the studies related to learning in chapter 2 of this thesis includes points and/or badges, which is mechanics suited to the two of four player types (Heeter, Magerko et al. 2009) mentioned earlier in this thesis. Player types that are considered according to the user group and the context the application is going to be used in. These studies show mixed results, and would be expected to differ from this thesis as well. However, points and badges will be included in this thesis, since it is said to be essential in a learning context to increase user motivation. Points are used as a highlight whenever a user is progressing through the application, and is perceived as a type of reward (Raymer 2011). Memrise, Fun Easy Learning Norwegian, Migranorsk and Duolingo use points. Since points seem to be consistent in both literature and practice, it will be used in the prototype as well. Users get the feeling of visible achievement and progress in learning. Duolingo is the only application that uses badges. McDaniel, Lindgren et al. (2012) talked about making badges visible and predictable for the users to appreciate them, and in this case they are displayed as a bronze, silver and gold medal of every exercise.

4.2.2 Research Question 2: Design for Gamification

Karl Kapp (2013) defines that there are two ways to apply gamification:

- *Structural Gamification* – The way this approach works, is that gamification mechanics are applied to an application without actually changing the content at all. It is only the mechanics that are added, and nothing else changes, except that the user may get points or badges for completing tasks that otherwise had no game elements associated with it. This type of gamification is focusing on giving users rewards to motivate them in the learning process.
- *Content Gamification* – This is the type of gamification that focuses on adjusting content to make it more like a video game. It is done by adding game-thinking and -elements. Game activities can be added to the content being taught.

Structural gamification is chosen in this case since it makes for an easier and more valid comparison of the two conditions. That means the user interface do not have to take gamification into account when being designed. The mechanics are merely added as a layer upon the already designed interface. The user interface can be more or less the same, where the gamification mechanics are the only factors that separate the conditions, since it is an important component in the experimental design. It is better for the sake of the experiment to make two conditions where only the gamification mechanics are differential factors. The gamification mechanics that have been chosen are also easily implemented into a non-game design.

Based on the interview findings, independence scored high as a frequent answer. For the users to be independent, the user interface has to be intuitive. Since usability testing and iteration of design are methods that are excluded in accordance with the ramifications of time in this project, it has to be in some sense similar to what learning center employees are currently using with their students. Therefore, the design was going to be based partially on Migranorsk, since that was the most popular software that there was access to. General design features reviewed in the state of the art section will also be included, since these are applications for mobile devices. These applications are also intuitive and with simple navigational features. The design features taken from the state of the art review is listed below:

- The user interface is maintained with as little clutter as possible, for limiting input and making an intuitive navigation.
- Tasks should be divided into modules, since this reduces cognitive effort and makes language learners learn incrementally (Raymer 2011).
- The type of tasks (Make a sentence, write the missing word, etc.) should be mixed in each topic, instead of letting the user choose the type of task. This way the number of clicks needed to start an exercise is reduced by one, reducing confusion, as well as letting the users get the variation in tasks they need.
- Tasks are designed to be repetitive, and can be performed more than once. The interface indicates that this is possible.
- Green and red is used for displaying correct or wrong answers (respectively) to exercises.

The first prototype will consist of two scenarios. Each scenario will have a gamified and non-gamified version. Both scenarios will focus on completing a language exercise. The scenarios will include different types of exercises to see if this may be a factor for satisfaction by the students. The main goal is to see if the scenarios with gamification give the user a greater degree of satisfaction. The navigation and differences in conditions of the prototypes can be seen in Appendix B.

4.2.3 Limitations in Design

As stated in "Common European Framework of Reference for Languages" (Modern Languages Division 2001), the primary language skills that should be learned are ***writing***, ***reading***, ***speaking*** and ***listening***. Mobile devices have the possibility to assist in learning all these skills, and a variety of the tasks should optimally be tested to find out if it has an impact on the student's satisfaction. However, as a prototype have limitations of functionality, the chosen tasks will exclude the auditive exercises, and only include to the visual ones. The chosen exercises are:

- Spelling of words/ a sentence by either writing from the mobile keypad or from a list of letters. This exercise is featured in all of the previously reviewed mobile applications.
- Select the missing word in a displayed sentence. This exercise is featured in both Babbel and Duolingo. Excluding the social aspect of the application is done knowingly, and because a function like that is impossible in the prototyping tool that is used.
- Another challenge is to make the interface dynamic enough to create the "feeling" of gamification. Gamification mechanics are dynamic elements, and would be performed best if made in a native application.

4.3 Experimental Research Results

The experimental research was conducted to either prove or reject the following hypothesis of this thesis: "*Satisfaction of mobile language learning is better perceived by users as opposed to an application not using gamification"*. To process the data that was collected, the data was put in a digital spreadsheet and imported to a software for statistical calculation called SPSS (Statistical Package for the Social Sciences). The software was used with a license provided by Gjøvik University College.

The experiment consisted of 14 participants (three female; 11 male; age 17 – 48) where nine were from Africa, four from Asia and one from Europe. All participants performed both conditions. The statistical method for analyzing the measured results was a *paired-samples t-test* (Lazar, Feng et al. 2010, p.75). The test is commonly used for a within-group experimental design, when comparing two conditions and there is one independent variable. The scores from both conditions – the gamified condition and the non-gamified condition – were individually calculated to a final SUS-score and then compared in the paired-samples t-test. The results of the questionnaire ended up with rejecting the hypothesis. There are no statistical significance in the analyzed result ($t = -1.095$, $p > 0.05$) that are suggesting otherwise (Figure 23). The entire output of the statistical data analysis can be seen in Appendix F.

Paired Samples Test

		Paired ...			
		95% Confidence Interval of the ...			
		Upper	t	df	Sig. (2-tailed)
Pair 1	nongame_score - game_score	2.259	-1.095	13	.293

Figure 23 - Figure of the analyzed results.

When participants were asked after the experiment about which condition they liked the most, five of 14 participants answered they preferred gamification because of the points they were collecting. The remainder of participants answered that they had no preference among the conditions and were equally pleased with both. The results would however not support the hypothesis made in this thesis, and will be discussed further in the next chapter.

5. Discussion and further work

5.1 Discussion of Results

This thesis' aim was to see if gamification mechanics had a positive satisfactory effect on learners of the Norwegian language. The prototype was designed by reviewing existing literature and applications on the subject, as well as using data from the interviews that were conducted both early in the project phase and in the pre-study. Participants of the user studies were mainly employees at the learning centers, but also students performing questionnaires from an earlier study. The prototype was evaluated by participants by answering a SUS questionnaire, and evaluated through a paired-samples t-test. The results based on testing statistical significance on the questionnaire scores rejected the hypothesis of this thesis.

The final experiment consisted of two groups that both were immigrants attending Norwegian courses. The first group was the class that was deemed the most proficient in Norwegian by the teachers at Gran Voksenopplæring (About to take the B1 test). The class that was participating the next day was just below the first group in language skill. It was however a noticeable difference in vocal understanding and performance. The study had to be conducted in Norwegian due to several ethnicities participating. The students in the first class were mostly very proficient in the Norwegian language, and they had no problems understanding the conversation and explanations they got. They may however have been reluctant to ask for explanations when answering the SUS to avoid feeling uncomfortable. The second group was overall less proficient in using the Norwegian language. Most of the students were progressing towards B1, but there were differences among the students within the class. The last group's combined score showed a clear difference as opposed to the first group.

In general, a user with a lot of prior experience will rate an application as more usable, hence the learning effect. It did however not show significance in this thesis, based on which condition the user started with. The random assignment was therefore of no consequence for the results.

As stated in the results of the experimental design, the result was not statistical significant to support the hypothesis. However, all students answered that they preferred either the gamified condition or that they had no preference at all. Neither age nor ethnicity had any impact on the result either. All participants said they found the application easy to use. That may have been a factor for the high score on the System Usability Scale. Most of the questions addressed learnability and complexity. If only these questions are taken in consideration, it can be argued for that a Wilcoxon signed ranks test (non-parametric test) can be used as an alternative of a paired-samples t-test, because a normal distribution cannot be assumed. If only answers expecting a high score on the topic of ease of use are compared between the conditions, the Wilcoxon signed ranks test will support the hypothesis in this thesis (Figure 24). It will however not be considered the proper result of this thesis, since the data in that case would have been in some way selected as was appropriate by the researcher.

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The median of differences between Nongame_ease and Game_ease equals 0.	Related-Samples Wilcoxon Signed Rank Test	.039	Reject the null hypothesis.

Asymptotic significances are displayed. The significance level is .05.

Figure 24 - Figure of result using Wilcoxon signed-rank test.

The only reason this method is mentioned, is that the results may show some signs of improved satisfaction that can be explored further.

5.2 Weakness of the Experimental Design.

The biggest barrier when performing the user test was that of language. All of the communication was done in Norwegian, and that meant explanations as well when the participants did not understand specific terms and words. Explanations may not have been sufficient, and it can be assumed that participants avoided further questioning to avoid feeling uncomfortable. A good example of this is when a participant asked if answering number five on the first question in the questionnaire was positive. The participant continued to answered number five even though it was only on odd-numbered questions where five was a positive answer. None of the participants mentioned any aspects they disliked about either scenario. It can therefore be reasonable to interpret that the questions were misunderstood by participants.

Results may have been affected by that none of the users had previously used applications to learn Norwegian, and had nothing to compare the tested application to. It seemed that only by having the opportunity to learn Norwegian through a mobile application made the students rate the application with a top score. Another explanation could be that the researcher was sitting too close to the participant, and somehow affected the participant by being present. This was however a dilemma, since the researcher had to be present in order to explain and conduct the experiment. The sample size was also somewhat lacking to make the result solid and generalizable. By doubling the amount of participants, the data collected may have ended up showing significance.

Thirdly, the measuring instrument could not be proven reliable enough for the purpose of this thesis. It can be shown in figure 25. If the user is satisfied with all aspects of the system, every odd number is supposed to generate a positive number, and all even numbers the opposite. By grouping answers to questions as variables in a rotated component matrix, it can be seen that this is not the distribution of answers that is anticipated.

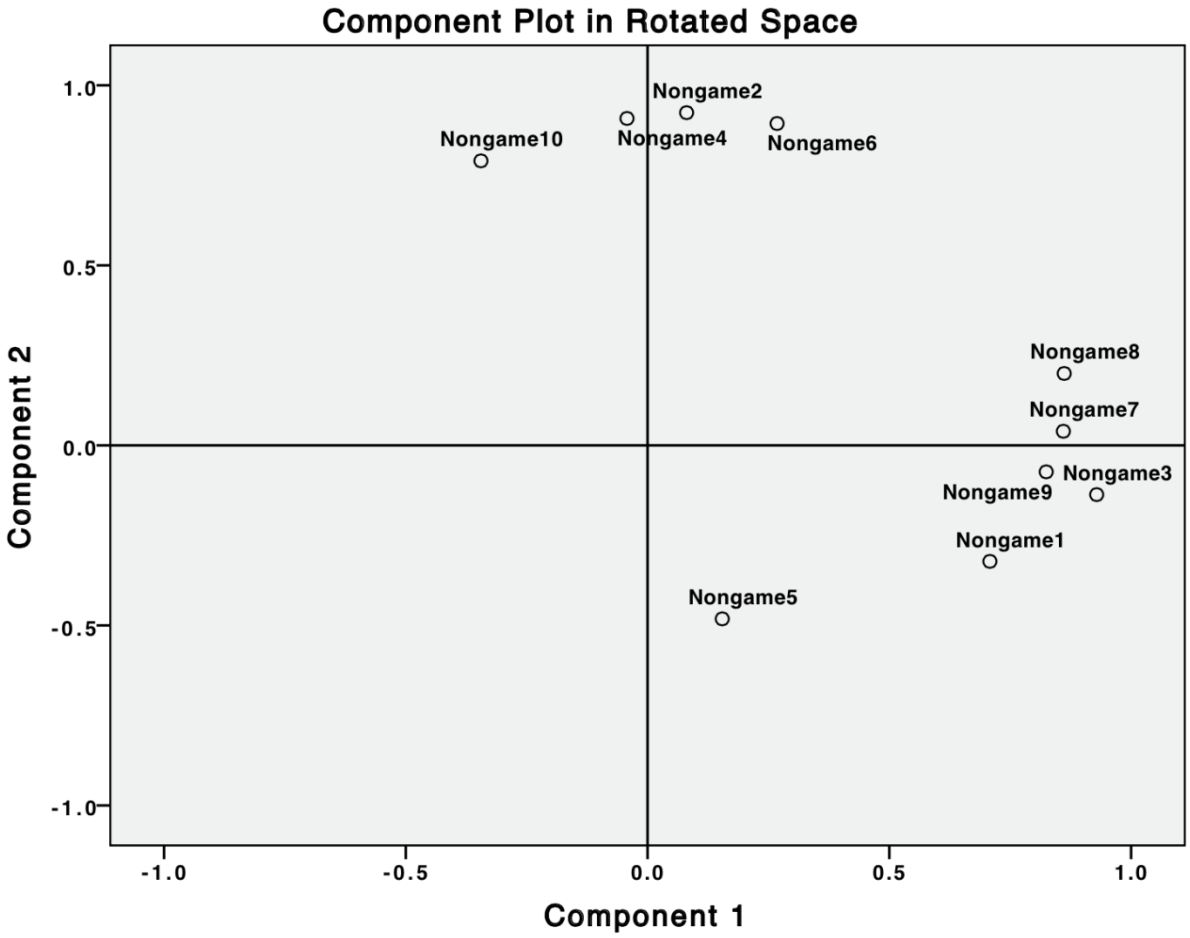


Figure 25 - Cluster of how questions were grouped.

6. Conclusion

The conclusion of this study is that even though the research done points to increased satisfaction of a gamified interface for mobile language learning, the results are not statistically significant to the extent to support the hypothesis. There may have been a weakness in the measuring instrument that has excluded satisfactory factors as esthetics and flow of interactions. Therefore, the way satisfaction was measured may not have justified the definition of the word.

6.1 Further Work

Further work on this study can involve a longitudinal study on the effects of gamification on a mobile device as *supplementary* learning. The frequency of usage can be measured to see if it corresponds with the results to Hanus and Fox's (2015) study of adding gamification in an analog setting. The language skills can also be tested both before and after with a control group. As this thesis is for a school semester, developing and performing a longitudinal study is impossible for the given time period. Performing a longitudinal study on the participants could have been interesting, as well as on a greater number of participants with a control group. For the users to better get an impression of the difference with and without the mechanics, the testing period should have lasted longer, and the participants should have been able to keep the prototype to collect the badges. The feeling of achievement may have manifested if the user would be able to keep the rewards, as stated by Lyytinen, Ronimus et al. (2007). The possibility of users keeping the application over a given time to experience the achievements could count as a stronger experimental design, something that is mentioned by Seaborn and Fels (2015) to be needed in the area of gamification.

Another longitudinal study can be done on the results of learning as Vesselinov and Grego (2012) discussed. They restricted their study to Spanish, but could not generalize the result for other languages. The reported satisfaction of the users was almost exclusively positive, and while the results of this thesis are not the same, they are pointing in the same direction. It was however two different instruments of measurement that was used in these studies. Seaborn and Fels (2015) - in their peer review of several research articles concerning gamification - concludes that

more research is needed of including both context of gamification in learning, demography of users and using stronger *experimental designs* – a research design that isolates factors in the area of psychology and computer science. Gamification based on research published is showing potential in several contexts. Further work consists of analyzing and mapping the contexts and how they respond to gamification.

It seems that for a better result for this thesis, satisfaction should be measured otherwise. Satisfaction should be analyzed and divided into factors affecting it, made into separate questions and answered individually by the participants. The study should longitudinal, since participants can have more time to explore the application. The short introduction and test of the application may have produced a novelty effect on the participants, and therefore generated a high SUS-score.

7. Bibliography

Apple (2014). "Appstore." from https://itunes.apple.com/no/genre/ios/id36?mt=8

Attali, Y. and M. Arieli-Attali (2015). "Gamification in assessment: Do points affect test performance?" Computers & Education **83**(0): 57-63.

Bartle, R. (1996). "Hearts, clubs, diamonds, spades: Players who suit MUDs." Journal of MUD research **1**(1): 19.

Bevan, N. (2006). "International standards for HCI." Encyclopedia of human computer interaction **362**.

Brooke, J. (1996). "SUS-A quick and dirty usability scale." Usability evaluation in industry **189**(194): 4-7.

Bureau, U. S. C. (2010). "Hispanic Heritage Month 2010: Sept. 15 – Oct. 15." Retrieved 10.12, 2014, from http://www.census.gov/newsroom/releases/archives/facts_for_features_special_editions/cb10-ff17.html.

Calisir, F. and F. Calisir (2004). "The relation of interface usability characteristics, perceived usefulness, and perceived ease of use to end-user satisfaction with enterprise resource planning (ERP) systems." Computers in Human Behavior **20**(4): 505-515.

Chen, C.-h. and K. Huang (2014). "The Effects of Response Modes and Cues on Language Learning, Cognitive Load and Self-Efficacy Beliefs in Web-based Learning." Journal of Educational Multimedia and Hypermedia **23**(2): 117-134.

Clark, R. C. and R. E. Mayer (2011). E-learning and the science of instruction: Proven guidelines for consumers and designers of multimedia learning, John Wiley & Sons.

de-Marcos, L., et al. (2014). "An empirical study comparing gamification and social networking on e-learning." Computers & Education **75**(0): 82-91.

Denny, P. (2013). The effect of virtual achievements on student engagement. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. Paris, France, ACM**:** 763-772.

Division, M. L. (2001). Common European Framework of Reference For Languages: Learning, Teaching, Assessment. Strasbourg, Cambridge University Press.

Domínguez, A., et al. (2013). "Gamifying learning experiences: Practical implications and outcomes." Computers & Education **63**(0): 380-392.

Google (2015). "Google Play." from https://play.google.com/store.

Hamari, J. (2015). "Do badges increase user activity? A field experiment on the effects of gamification." Computers in Human Behavior(0).

Hamari, J. and J. Koivisto (2013). Social Motivations To Use Gamification: An Empirical Study Of Gamifying Exercise. ECIS.

Hamari, J., et al. (2014). Does Gamification Work? -- A Literature Review of Empirical Studies on Gamification. System Sciences (HICSS), 2014 47th Hawaii International Conference on.

Hanus, M. D. and J. Fox (2015). "Assessing the effects of gamification in the classroom: A longitudinal study on intrinsic motivation, social comparison, satisfaction, effort, and academic performance." Computers & Education **80**(0): 152-161.

Hashemi, M., et al. (2011). "What is mobile learning? Challenges and capabilities." Procedia-Social and Behavioral Sciences **30**: 2477-2481.

Heeter, C., et al. (2009). "Game design and the challenge-avoiding, self-validator player type." International Journal of Gaming and Computer-Mediated Simulations (IJGCMS) **1**(3): 53-67.

Holmes, B. and J. Gardner (2006). E-learning: Concepts and practice, Sage.

Jordan, P. W., et al. (1996). Usability evaluation in industry, CRC Press.

Jovanovic, D. and M. Matejevic (2014). "Relationship between Rewards and Intrinsic Motivation for Learning–Researches Review." Procedia-Social and Behavioral Sciences **149**: 456-460.

Kapp, K. (2013). "Two Types of #Gamification." from http://karlkapp.com/two-types-of-gamification/.

Kukulska-Hulme, A. (2009). "Will Mobile Learning Change Language Learning?" ReCALL **21**(2): 157-165.

Lazar, J., et al. (2010). Research methods in human-computer interaction. Chichester, John Wiley.

Li, W., et al. (2012). GamiCAD: a gamified tutorial system for first time autocad users. Proceedings of the 25th annual ACM symposium on User interface software and technology. Cambridge, Massachusetts, USA, ACM**:** 103-112.

Lyytinen, H., et al. (2007). "Early identification of dyslexia and the use of computer game-based practice to support reading acquisition." Nordic Psychology **59**(2): 109.

McDaniel, R., et al. (2012). Using badges for shaping interactions in online learning environments. Professional Communication Conference (IPCC), 2012 IEEE International, IEEE.

Miangah, T. M. and A. Nezarat (2012). "Mobile-assisted language learning." International Journal of distributed and Parallel System **3**: 309-319.

Myhre, M. G. (2014). A User Study of Language Learning on Electronical Devices. IMT 4882 Specialization Course 2, Gjøvik University College.

Nielsen, J. (1993). Usability engineering. Boston, Mass, Academic Press.

NSD (2012). "Personvernombudet for Forskning." Retrieved 16.12, 2014, from http://www.nsd.uib.no/personvern.

Ozuorcun, N. C. and F. Tabak (2012). "Is M-learning Versus E-learning or are They Supporting Each Other?" Procedia - Social and Behavioral Sciences **46**(0): 299-305.

Pereira, O. R. E. and J. J. P. C. Rodrigues (2014). "Survey and analysis of current mobile learning applications and technologies.(Report)(Author abstract)." ACM Computing Surveys **46**(2): 27.

Perry, B. (2015). "Gamifying French Language Learning: A Case Study Examining a Quest-based, Augmented Reality Mobile Learning-tool." Procedia - Social and Behavioral Sciences **174**(0): 2308-2315.

Raymer, R. (2011). "Gamification: Using Game Mechanics to Enhance eLearning." eLearn **2011**(9).

Rubin, J. (2008). Handbook of usability testing : how to plan, design, and conduct effective tests. Indianapolis, Wiley.

Ryan, R. M., et al. (2000). "Self-Determination Theory and the Facilitation of Intrinsic Motivation, Social Development, and Well-Being." American Psychologist **55**(1): 68-78.

Saffer, D. (2010). Designing for interaction : creating innovative applications and devices. Berkeley, Calif, New Riders.

Seaborn, K. and D. I. Fels (2015). "Gamification in theory and action: A survey." International Journal of Human-Computer Studies **74**(0): 14-31.

Seaborn, K. and D. I. Fels (2015). "Gamification in theory and action: A survey." International Journal of Human-Computer Studies **74**: 14-31.

Segaran, K., et al. (2014). "Usability and User Satisfaction of 3D Talking-head Mobile Assisted Language Learning (MALL) App for Non-native Speakers." Procedia - Social and Behavioral Sciences **131**(0): 4-10.

Seong, D. S. K. (2006). Usability guidelines for designing mobile learning portals. Mobility '06, ACM**:** 25-es.

Sharp, H., et al. (2015). Interaction design : beyond human-computer interaction. Chichester, Wiley.

Sharples, M. (2006). "Big issues in mobile learning."

Sharples, M. and J. Roschelle (2010). "Guest editorial: Special section on mobile and ubiquitous technologies for learning." IEEE Transactions on Learning Technologies **3**(1): 4-6.

Stockwell, G. and S. Sotillo (2011). "Call for Papers for Special Issue of LLT. Theme: Mobile Language Learning." Language Learning & Technology **15**(3): 130.

Su, C.-H. and C.-H. Cheng (2013). "A Mobile Game-based Insect Learning System for Improving the Learning Achievements." Procedia - Social and Behavioral Sciences **103**(0): 42-50.

Trentin, G. and M. Repetto (2013). Using Network and Mobile Technology to Bridge Formal and Informal Learning. Oxford, Chandos Publishing.

Vesselinov, R. and J. Grego (2012). "Duolingo effectiveness study." City University of New York, USA.

Werbach, K. and D. Hunter (2012). For the win: How game thinking can revolutionize your business, Wharton Digital Press.

Zichermann, G. and C. Cunningham (2011). Gamification by Design : implementing game mechanics in web and mobile apps. Sabastobol , CA, O'Reilly Media.

Appendix A – Interview Form

Intervjuguide

Kan du fortelle kort om din typiske arbeidsdag?
Can you tell me briefly about a typical day at work?

Hvor lenge har du undervist her på Læringssenteret?
How long have you been teaching here at Læringssenteret?

Hvilke nivåer i språkferdighet underviser du i?
What levels of language proficiency do you teach?

Hvilke læringsstiler tar du i bruk i din undervisning? (Visuell, auditiv, verbal, fysisk, logisk, sosial, ensom?)
Which learning styles do you use while teaching?

Med tanke på din klasse, hvilken læringsstil tror du er mest effektiv da du underviser i et klasserom?
With your class in mind, which learning styles do you think are most effective when you teach in a classroom?

Med tanke på din klasse, hvilken læringsstil tror du er mest effektiv da elevene bruker elektroniske opplæringsprogram?
With your class in mind, what learning styles do you think are most effective when the students use digital language applications?

Jeg ser i læreplanen at hver students digitale basisferdigheter blir kartlagt. Hvor ligger gjennomsnittlig dette nivået på?
I see in the curriculum that each student's digital skills are mapped. What is this average level of this skill?

Hvordan foregår opplæringen i digitale ferdigheter?
How is the training in digital skills executed?

Brukes det teknologi i opplæring annet enn gjennom elektroniske opplæringsprogram? (Filmer, nettsider med ordbok osv.)
Is there any use of technology in education other than through digital language applications? (Movies, websites with dictionaries, etc..).

Hvilke elektroniske opplæringsprogram har du erfaring med?
Which digital language applications do you have experience with?

(Om flere) Hvilke elektroniske opplæringsprogram har du best erfaring med, og hvorfor?
(If more than one) Which digital language applications do you have the best experience with and why?

Hva er du mest fornøyd med i de elektroniske opplæringsprogrammene som blir brukt?
What are you most satisfied with in the digital language applications that are used?

Finnes det frustrasjonsmomenter i disse/ dette programmet hos enten lærere eller elever? (Eventuelt hvilke?)
Are there moments of frustration in these / this digital language applications by either teachers or students? (If yes, which?)

Hva har du inntrykk av at er den viktigste grunnen til at elektronisk opplæringsprogram blir brukt til opplæring?
What is your impression of what is the most important reason digital language applications are used in learning?

Hva synes du selv er fordelene ved å bruke elektroniske opplæringsprogram i språkopplæringen?
Which advantages do you see in using digital language applications?

Hva tror du er ulempen med å inkludere teknologi i opplæringen?
What disadvantages do you think technology in education entails?

Hvordan tror du interessen til elevene generelt er mot opplæring ved hjelp av teknologi?
How do you think the interest of students in general are with education through technology?

Føler du at noe mangler/ kunne blitt gjort annerledes med tanke på teknologien som blir brukt i opplæringsprogrammet?
Do you feel that something is missing / could have been done differently regarding the tehnology in the education programme?

Har du inntrykk av at de fleste elevene eier en mobiltelefon?
Do you think most of the students own a mobile phone?

Har du noen erfaring med språklæring på mobil?
Do you have any experience with language learning on mobile?

(Om ja) I tilfelle hvilke applikasjoner?
(If yes) Which applications?

(Om ja) Jeg ser i læreplanen at deltakere må i samarbeid med lærere utvikle en læringsstrategi og tildels kunne styre sin egen læring. Tror du dette er noe som kunne vært innført i en læringsstrategi?
(If yes) I see in the curriculum that participants in cooperation with teachers need to develop a learning strategy and partly be able to control their own education. Do you think this is something that could be implemented in a learning strategy?

Appendix B – Prototypes

Figure 1 - Choose Subject

Figure 2 - Choose exercise

Figure 3 - Start exercise

Figure 4 - Introduction for words 1

Figure 5 - Introduction for words 2

Figure 6 - Introduction for words 3

Figure 7 - Exercise type 1, task number 1

Figure 8 - Exercise type 1, task number 2

Figure 9 - Exercise type 1, task number 3

Figure 10 - First type of exercise completed

Figure 11 - Exercise type 2, task number 1

Figure 12 - Exercise type 2, task number 2

Figure 13 - Exercise type 2, task number 3

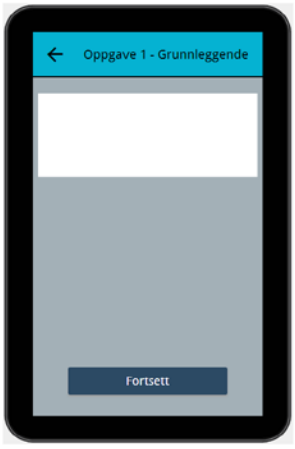


Figure 14 - Second type of exercise completed

Figure 15 - Exercise type 3, task number 1 in progress

Figure 16 - Exercise type 3, task number 1 finished

Figure 17 - Exercise type 3, task number 2 in progress

Figure 18 - Exercise type 3, task number 2 finished

Figure 19 - End-screen of "Oppgave 1" non-gamified

Figure 20 - End-screen of "Oppgave 1" gamified

Gamified Condition

The feedback on the gamified condition shows "+1 Poeng" while the buttons disappear when answering correct. The answer is also displayed in the chat-bubble at the top of the screen.

The feedback on the gamified condition shows the button pressed blinking red when answering incorrect. The answer is not displayed in the chat-bubble. Instead there is a red line where the answer should have been.

Non-gamified Condition

The feedback on the non-gamified condition shows the button pressed blinking green when answering correct. This is consistent for all exercises except the first, where the correct word will be filled in a chat-bubble.

The feedback on the non-gamified condition shows the button pressed blinking red when answering incorrect.

The ending screen of each exercise shows the total acquired points collected by answering correct. Based on the users performance, a medal is awarded as a badge that the user will keep.

The ending screen of each exercise shows a simple notification that the exercise is completed and instructs the user that the button at the bottom of the screen takes him/her back to the main menu

Appendix C – Consent Form

Samtykkeskjema

Bakgrunn

Mobiltelefoner og applikasjoner er i dag brukt av mange mennesker til mange formål. Et av målene applikasjoner oppfyller i dag er språklæring. Grunnen til at dette er populært på mobiltelefoner er at de er enheter man kan ta med seg overalt. Det gjør at man også kan ta med seg læringen overalt, og utføre den når som helst. Om oppgavene er korte og raskt gjennomførbare er dette noe som kan øke effektiviteten av språklæring. Det som også er kjent ved digital og selvstendig læring, er innføring av spillelementer(spillifisering). Dette er sagt at kan øke motivasjonen hos brukerne, noe som har satt i gang dette prosjektet.

Eksperimentet

Eksperimentet går ut på at to scenarioer skal testes for å se om spillifisering har et utslag på hvor tilfreds deltakerne opplever applikasjonen. Målingen av tilfredshet gjøres ved at deltakeren fyller ut et skjema etter testing. I tillegg vil noen spørsmål bli stilt til slutt slik at årsaken til tilfredsheten kan analyseres grundigere.

Praktisk informasjon

Eksperimentet består av to scenarioer, to skjema (SUS) som skal fylles ut etter hvert scenario, og et raskt intervju til slutt. Det er forventet at det vil foregå i ca. 20 minutter fra start til slutt. Videoopptak gjøres under utførelsen av hvert scenario, og skal bli brukt kun for min egen del til datainnsamling. Ingen av videoene vil bli publisert, og hver deltaker forblir anonym etter prosjektets slutt. Det er viktig å nevne at du kan trekke deg til enhver tid i løpet av eksperimentet.

Ved å signere på dette skjemaet har deltakeren forstått all relevant informasjon om utføring av eksperimentet, og bakgrunnen til at det gjennomføres.

_____ _____

Deltakers signatur Dato/Sted

_____

Fasilitators signatur

Appendix D – Manuscript

Manus Eksperiment SmartNorsk

Gjør klar notat for error-notering

“Hei, og takk for at du vil delta på denne brukertesten til min masteroppgave. Den går ut på at jeg skal lage en mobilapplikasjon for å lære norsk. Jeg vil bare si med en gang at ingen personlig informasjon om deg vil bli publisert, og du er helt anonym gjennom hele oppgaven. Jeg vil gjerne at du leser og skriver under på et samtykkeskjema før vi begynner, slik at vi er enige om at informasjonen vi samler kan bli brukt til oppgaven min.”

Gi ut samtykkeskjema og forklar om nødvendig

“Da kan vi begynne. Jeg vil gjerne at du tenker høyt mens du gjennomfører oppgaven. Si gjerne hvorfor du gjør det du gjør underveis.”

Velg ut scenario 1 eller 2 på mobilen

“Jeg vil gjerne at du skal finne oppgave 1 i Grunnleggende og utføre den”

Bruker utfører scenario

Takk, nå vil jeg gjerne at du svarer på denne.

Gi SUS til deltakeren

“Kjempebra! Nå vil jeg gjerne at du gjør det samme en gang til på denne applikasjonen”

Repeter for Scenario 2

“Da vil jeg bare stille deg noen spørsmål til applikasjonen du nettopp fikk prøve. Denne applikasjonen er under utvikling, og du må være så ærlig du bare kan når du svarer slik at jeg kan forbedre applikasjonen. Hvis det er noe du ikke forstod, er det noe galt med applikasjonen, ikke deg.

1. Hvilken versjon likte du best? (gamified eller non-gamified?)
2. Var applikasjonen lett å bruke?
3. Var applikasjonen lett å forstå selv om alt var på norsk?
4. Hva syntes du om språkoppgaven du fikk? (Lett? Vanskelig?)
5. Var det noe du du synes var vanskelig å forstå på applikasjonen? (Evt. Hva?) Var det noe du likte ved applikasjonen?
6. Eier du en smarttelefon? (Evt. Hvor ofte bruker du den, og hva bruker du den til?)” Ville du brukt en lignende applikasjon til å øve på norsk?

“Tusen takk for deltakelsen! Da er du ferdig (og kan hente neste person)”

Appendix E – System Usability Scale

Spørreundersøkelse

Sett ring rundt det svaret du føler passer din opplevelse av vårt system. Husk at ved å være 100% ærlig hjelper du meg mest.

1. Jeg kunne tenke meg å bruke denne applikasjonen ofte.

1 2 3 4 5

Svært uenig *Svært enig*

2. Jeg synes applikasjonen inneholdt unødvendig/kompleks informasjon.

1 2 3 4 5

Svært uenig *Svært enig*

3. Jeg synes applikasjonen var lett å bruke.

1 2 3 4 5

Svært uenig *Svært enig*

4. Jeg tror jeg ville hatt behov for støtte fra en ekspert for å bruke denne applikasjonen.

1 2 3 4 5

Svært uenig *Svært enig*

5. Jeg synes de forskjellige oppgavene i applikasjonen er godt integrert.

1 2 3 4 5

Svært uenig *Svært enig*

6. Jeg synes applikasjonen var lite konsekvent (Innholder mange forskjellige elementer slik at den er vanskelig å skjønne).

1 2 3 4 5

Svært uenig *Svært enig*

7. Jeg kan tenke meg at de fleste kan lære seg å bruke applikasjonen veldig raskt.

1 2 3 4 5

Svært uenig *Svært enig*

8. Jeg synes applikasjonen var veldig vanskelig å bruke.

1 2 3 4 5

Svært uenig *Svært enig*

9. Jeg følte meg veldig sikker på hvordan jeg skulle bruke applikasjonen.

1 2 3 4 5

Svært uenig *Svært enig*

10. Jeg måtte lære meg mye før jeg kunne starte å bruke applikasjonen.

1 2 3 4 5

Svært uenig *Svært enig*

Appendix F – Statistical Analysis Output

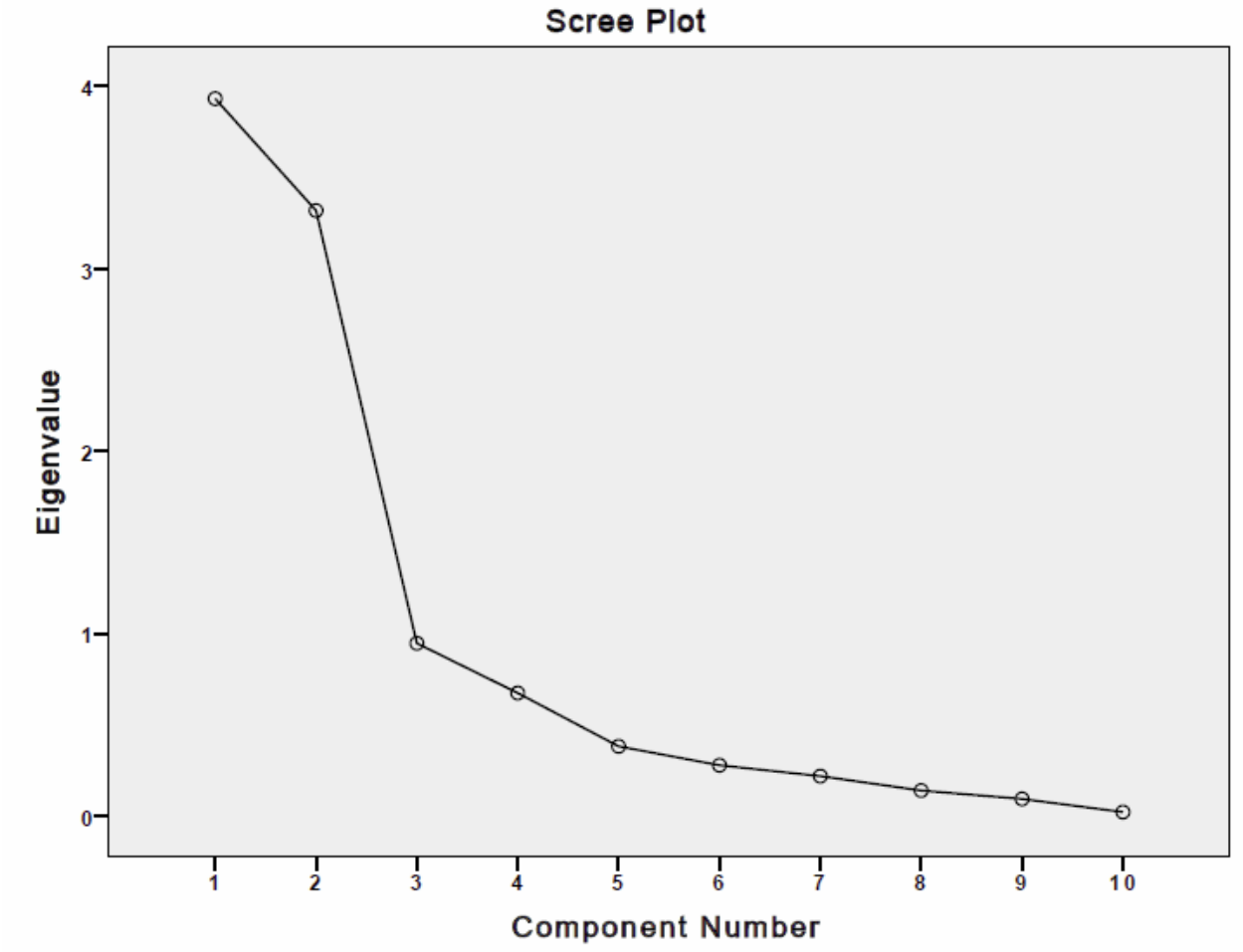
Factor Analysis

Notes

Output Created		13-MAY-2015 12:50:49
Comments		
Input	Data	/Users/Morten/Desktop/RawScoresSUS.sav
	Active Dataset	DataSet1
	Filter	<none>
	Weight	<none>
	Split File	<none>
	N of Rows in Working Data File	14
Missing Value Handling	Definition of Missing	MISSING=EXCLUDE: User-defined missing values are treated as missing.
	Cases Used	LISTWISE: Statistics are based on cases with no missing values for any variable used.
Syntax		FACTOR /VARIABLES Nongame1 Nongame2 Nongame3 Nongame4 Nongame5 Nongame6 Nongame7 Nongame8 Nongame9 Nongame10 /MISSING LISTWISE /ANALYSIS Nongame1 Nongame2 Nongame3 Nongame4 Nongame5 Nongame6 Nongame7 Nongame8 Nongame9 Nongame10 /PRINT UNIVARIATE ROTATION /PLOT EIGEN ROTATION /CRITERIA MINEIGEN(1) ITERATE(25) /EXTRACTION PC /CRITERIA ITERATE(25) /ROTATION VARIMAX /METHOD=CORRELATION.
Resources	Processor Time	00:00:01.23
	Elapsed Time	00:00:01.00
	Maximum Memory Required	13688 (13.367K) bytes

Descriptive Statistics

	Mean	Std. Deviation	Analysis N
Nongame1	3.5714	.75593	14
Nongame2	3.1429	1.56191	14
Nongame3	3.5714	.85163	14
Nongame4	2.8571	1.61041	14
Nongame5	3.5714	.64621	14
Nongame6	2.7857	1.52812	14
Nongame7	3.5714	.85163	14
Nongame8	3.7143	.61125	14
Nongame9	3.5000	.65044	14
Nongame10	2.2857	1.63747	14



Component Matrix[a]

a. 2 components extracted.

Rotated Component Matrix[a]

	Component	
	1	2
Nongame1	.708	-.322
Nongame2	.081	.924
Nongame3	.929	-.137
Nongame4	-.043	.908
Nongame5	.155	-.482
Nongame6	.268	.894
Nongame7	.860	.039
Nongame8	.862	.200
Nongame9	.825	-.073
Nongame10	-.345	.790

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.

a. Rotation converged in 3 iterations.

Total Variance Explained

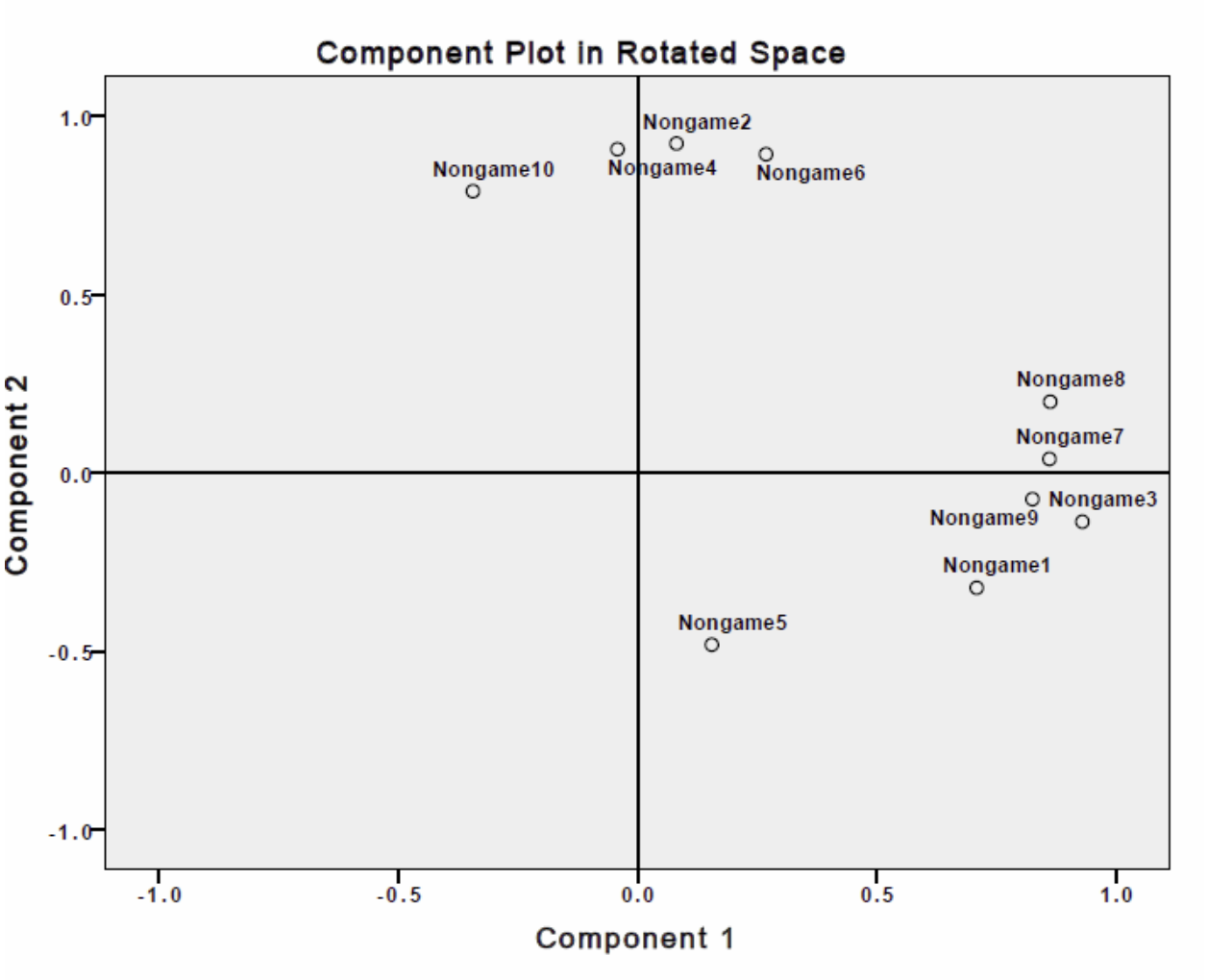
	Rotation Sums of Squared Loadings		
Component	Total	% of Variance	Cumulative %
1	3.749	37.486	37.486
2	3.503	35.032	72.518

Extraction Method: Principal Component Analysis.

Component Transformation Matrix

Component	1	2
1	.837	-.548
2	.548	.837

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.


Component Plot in Rotated Space
Component 2
Component 1
1.0
0.5
0.0
-0.5
-1.0
-1.0
-0.5
0.0
0.5
1.0
Nongame2
Nongame10
Nongame4
Nongame6
Nongame8
Nongame7
Nongame3
Nongame9
Nongame1
Nongame5

Paired sample t-test

T-Test

Notes

Output Created		12-MAY-2015 12:27:25
Comments		
Input	Active Dataset	DataSet0
	Filter	<none>
	Weight	<none>
	Split File	<none>
	N of Rows in Working Data File	14
Missing Value Handling	Definition of Missing	User defined missing values are treated as missing.
	Cases Used	Statistics for each analysis are based on the cases with no missing or out-of-range data for any variable in the analysis.
Syntax		T-TEST PAIRS=nongame_score WITH game_score (PAIRED) /CRITERIA=CI(.9500) /MISSING=ANALYSIS.
Resources	Processor Time	00:00:00.01
	Elapsed Time	00:00:00.00

[DataSet0]

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	nongame_score	81.43	14	14.991	4.006
	game_score	83.75	14	13.037	3.484

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	nongame_score & game_score	14	.849	.000

Paired Samples Test

		Paired Differences			
		Mean	Std. Deviation	Std. Error Mean	95% Confidence .. Lower
Pair 1	nongame_score - game_score	-2.321	7.934	2.120	-6.902

Paired Samples Test

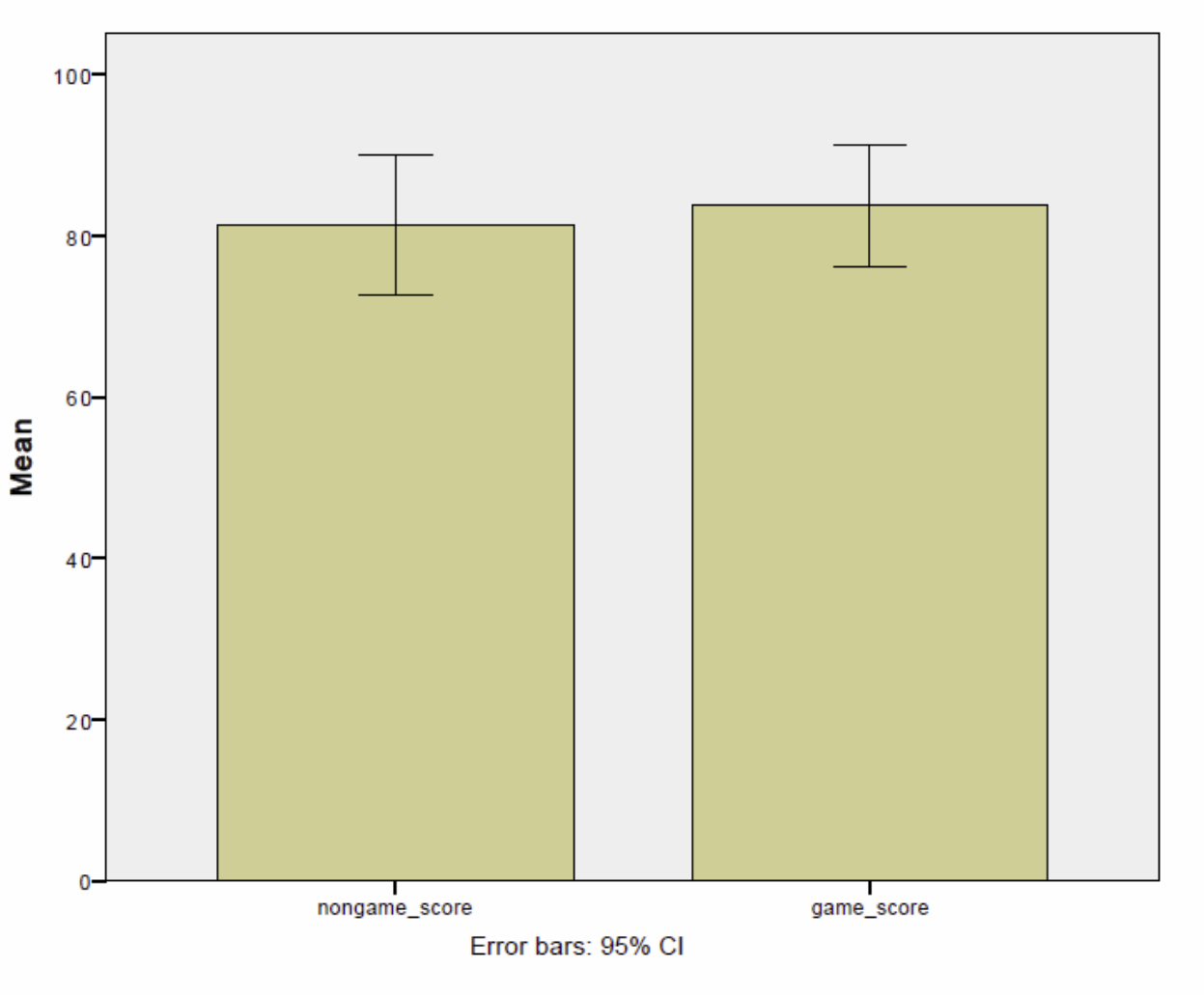
		Paired ...			
		95% Confidence Interval of the ... Upper	t	df	Sig. (2-tailed)
Pair 1	nongame_score - game_score	2.259	-1.095	13	.293

```
GRAPH
  /BAR(SIMPLE)=MEAN(nongame_score) MEAN(game_score)
  /MISSING=LISTWISE
  /INTERVAL CI(95.0).
```

Graph

Notes

Output Created		12-MAY-2015 12:36:29
Comments		
Input	Active Dataset	DataSet0
	Filter	<none>
	Weight	<none>
	Split File	<none>
	N of Rows in Working Data File	14
Syntax		GRAPH /BAR(SIMPLE)=MEAN (nongame_score) MEAN (game_score) /MISSING=LISTWISE /INTERVAL CI(95.0).
Resources	Processor Time	00:00:00.52
	Elapsed Time	00:00:00.00


100
80
60
40
20
0
Mean
nongame_score
game_score
Error bars: 95% CI