IPP report
Gamifying Spaced Repetition Software

s1700260

Abstract

Spaced repetition is a technique aimed to help students learn new concepts. Gamification is a method that adds game principles and elements to existing software to increase user engagement. The present report presents a plan to design, implement, and evaluate a gamified version of an existing spaced repetition software. The proposed solution integrates the spaced repetition application with an existing game such as users will be encouraged to use the spaced repetition features in order to acquire valuable resources to use in the game. An evaluation methodology, a set of scheduled tasks, and a risk management strategy are presented.

1. Introduction

Learning requires a considerable amount of mental effort. The amount of the mental effort expended in the learning process greatly depends on the perception of learners about the source of knowledge and the context of learning (Salomon, 1983). It follows that under conditions where people perceive the source of knowledge in a positive manner, the learning process requires less mental effort, therefore, it is easier, more pleasant, and provides better results (Salomon, 1984).

Traditional learning schemes have tried to include elements from new technologies with the objective of facilitating the learning process. One of such approaches is known as serious games (Michael & Chen, 2005). This alternative aims to take advantage of the pedagogical value of fun, competition and leisure to accomplish that objective. This technique has demonstrated to provide benefits in the acquisition of new knowledge and skills (Graafland et al., 2012).

Even though the proven benefits of serious games, it is narrowed to specific fields and contexts. Moreover, the considerable number of resources necessitated to develop such type of solutions makes it impractical to implement it in broader contexts. Under such circumstances, other alternatives have emerged to leverage the benefits of leisure and entertainment, but including characteristics of flexibility and adaptation. One of such approaches is known as gamification.

The main objective of gamification is to increase the user engagement of existing products, services or processes by including game elements and principles to make them more appealing. The range of scenarios of usage of this technique is wider and includes organizational productivity (Zichermann & Cunningham, 2011), physical exercise (Hamari & Koivisto, 2015) and learning (Hamari et al., 2016).

In the learning context, gamification has proved to be effective for different target audiences including primary school students (Boticki et al., 2015), undergraduate students (Slish et al., 2015), and the general public (DiSalvo, 2014). The types of learning products, platforms, and services that have implemented gamification are wide including mobile applications (Su & Cheng, 2015), websites (Gené et al., 2014) and desktop software (Cheong et al., 2013). These conditions ratify the flexible nature of gamification and its ability to be adapted to different environments.

Before the raise of technological techniques aimed to leverage the benefits of good perceptions of learning contexts and sources of knowledge, there were other alternatives aimed to facilitate the acquisition of new knowledge. One of these techniques was oriented to exploit the spacing effect phenomenon (Hintzman, 1974) to increase the capacity of retention of learners about specific contents. This technique is known as spaced repetition.

One of the foundations of spaced repetition is the acquisition of new content by recurrent revisions in a series of short learning sessions scattered over fixed or variable intervals of time. These conditions increase the capacity of retention compared to the acquirement of knowledge in a single massive learning session. Thus, the difficulty of remembering new facts, concepts or definitions is diminished and mental processes like recognition and recall are boosted.

Originally, spaced repetition is implemented with flashcards. Each one of them contains one or multiple related concepts. The flashcards are grouped based on how well the learner remembers their content. Then, the group of flashcards with more challenging content is presented to the learner more frequently. Flashcards move from one group to another as the learner keeps progressing in each session.

Spaced repetition has also taken advantage of new technologies. Thus, different pieces of software have been developed to make it more accessible through different platforms and interfaces. Each of these implementations has its own characteristics depending on the target audience. Some solutions are focused on topics like learning new languages

(von Ahn Luis, 2018), whereas others have broader and more general scopes (Marcus, 2018) and (Damien, 2018).

The current document presents a proposal to gamify an existing mobile application that implements spaced repetition. The ultimate objective is to create a more appealing version to increase the user engagement aimed to improve the effectiveness of spaced repetition. The following sections describe the design of the solution, the methodology for evaluation, the execution plan, and the risk management strategy.

2. Design

The usual way to implement gamification into existing solutions is to append game principles and elements throughout the application. An example of this approach is the addition of a leaderboard of the users based on points earned while using the application. Such leaderboard adds another motivation to use the application in order to reach the top of the classification. Overall, this way of implementing gamification means that the entire user experience can be converted into a game context.

For the current proposal, the way to implement gamification is different to the traditional one. Rather than appending game elements into the spaced repetition software, the project aims to fuse it with an existing game. Thus, two software components will be integrated as a single application. This condition poses a complexity in setting the balance of the main context for the final solution. Based on this, there are two alternatives to implement the application.

The first option requires the spaced repetition component to be the main activity of the final application. Such requirement means that using spaced repetition and playing the game would occur in interleaved stages as seen in Figure 1. The implementation of this option would be straightforward, however, a potential problem would be the lack of interconnection between the game and the spaced repetition component. Thus, each component might be interpreted as an interruption of the other.

For the second alternative, the game would play the main role. Evidently, a similar situation of interruptions might occur. However, that potential issue would be avoided by creating a true connection between the game and the spaced repetition component. The level of complexity of the implementation would be higher compared to the first option, but the potential benefits would be higher as well. The details of this approach along with the descriptions of the spaced repetition software and the game are described in the following subsections.

Figure 1. Flow of spaced repetition and game components with no real connection between them.

2.1. Integration of spaced repetition and an existing game

As described in the introduction section (1), when the source of knowledge is perceived in a positive way, the learning process requires less mental effort and it is more pleasant. Based on this idea, it is safe to assume that playing a game is perceived more positively than executing spaced repetition. For this reason, the main context of the final application will be the game component, and the spaced repetition component will complement it.

The final application will integrate the spaced repetition and game components as a single piece of software. Therefore, it is necessary to implement a connection between both components. As mentioned previously, one of the objectives of such connection is to avoid that the user interprets the components as interruptions of each other. However, the main goal is to establish a relationship such the output of using one component can be seen as a valuable resource to be used in the other component as seen in Figure 2.

Since the main context of the final application will be the game, the spaced repetition component needs to be connected in a way such the users find worth using that component. One alternative is to reward the users when using the spaced repetition component. Then the reward can be used during the game. Such rewards can be implemented as points to reach higher positions in a leaderboard, or as resources that can be used to ease the game.

A similar scheme was proposed in previous work (Johnson et al., 2012). There, the flow of the final application was split into two phases: points earning and game playing. During the first phase, the participants obtained points in two stages. In the first stage, the participants reviewed content using spaced repetition. The second stage consisted of a quiz where points were granted based on the number of correct answers. Then, those points could be used during the game.

There are other mechanisms that can be used to reward users and motivate them to use the spaced repetition component. However, for simplicity and ease of implementation, only the scheme of points earning will be taken into consideration. Therefore, the usage of the application will be split into two stages: points earning and game playing. The main difference with the scheme in (Johnson et al., 2012) is the flexibility of executing the points earning stage.

In the original scheme, the flow of the stages was fixed and the participants were able to earn points only at the beginning of each game session. For the current design, that flow is still valid and the points earning stage will be mandatory. However, the participants will also have a single extra opportunity to earn points. They will be able to execute the spaced repetition component at any time while playing the game. The application might suggest that option to the participants, especially during challenging levels.

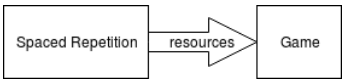

Figure 2. Flow of spaced repetition and game components with connection via resources provided from one component to the other.

2.2. The spaced repetition component: AnkiDroid

Among all the available spaced repetition software only a few of them are open source. The most relevant option is Anki (Damien, 2018) which is distributed under the GNU Affero General Public License (AGPLv3) for most of its platform versions except iOS. Therefore, it is possible to modify the existing source code and adapt it as the convenience.

Anki makes spaced repetition more accessible to a wider audience by providing web, desktop, and mobile interfaces. Anki also provides a group of characteristics including different types of content in the flashcards (text, images, sounds), a multitude of decks compiled for several languages and various topics, and statistics of usage. It also provides features to customize the topics to learn.

The Anki version for Android devices is known as AnkiDroid. It provides an intuitive user interface that follows the best practices for mobile development. However, the lack of elements that motivate its usage might make it boring for some users which makes it a good fit for the current project.

Figure 3. Example of AnkiDroid user interface. Image taken from the Google Play page of the application.

Spaced repetition in Anki is implemented using the SuperMemo 2 (SM-2) algorithm (**?**). The algorithm defines the next date to review an item depending on the current answer as seen in expression 1 where y is the easiness, and x is the consecutive correct answers for that item. Easiness is a number greater than or equal to 1.3 which represents how easy to learn is an item, being 1.3 the most difficult.

$$\text{Next date} = \begin{cases} now + 6y^{x-1}\ days & \text{if correct} \\ now + 1\ day & \text{if incorrect} \end{cases} \quad (1)$$

2.3. The game: 2048

Since the selected spaced repetition component is implemented for Android devices, it is necessary to choose a game for the same platform. Among the huge number of games available in Android, just a few of them are open source. The selected game for the current project is the popular 2048 (Uberspot, 2017). It is distributed under the MIT License, which is compatible with other licenses including AGPLv3. Therefore, it is possible to modify the code as the convenience as well.

2048 was first developed for web environments, but ever since it has been ported to many other platforms. It is an sliding puzzle which ultimate objective is to merge numbered blocks until create a tile with the number 2048. Blocks are distributed over a grid of cells of variable size, but the default number of cells is 4x4. At the start of the game, two randomly positioned tiles holding the number 2 are displayed in the grid as seen in Figure 4.

A player has to slide the blocks horizontally or vertically. The blocks are moved in the chosen direction and they are stopped by the edges or when colliding with other blocks. If two blocks holding the same numbers collide, they are merged into a single block holding the sum of the values of the previous blocks. In every turn a new block holding a power of two number is randomly positioned in the grid, being the power an integer between 1 and 10 as seen in expression (2);

$$\textit{Number in tile} = 2^x,\ x \in \{1, 2, \ldots, 10\} \quad (2)$$

The game ends if one of the following states is reached:

1. A tile holding the number 2048 is created.
2. The grid is full of blocks, none of them holding the number 2048.

The first state implies that the user has won. Nonetheless, the user can continue playing to create tiles with numbers higher than 2048. The second states indicates that the user has lost. Therefore, a new session has to be initiated to play again.

3. Evaluation

The evaluation process requires to measure the variation of user engagement in the gamified version of the spaced

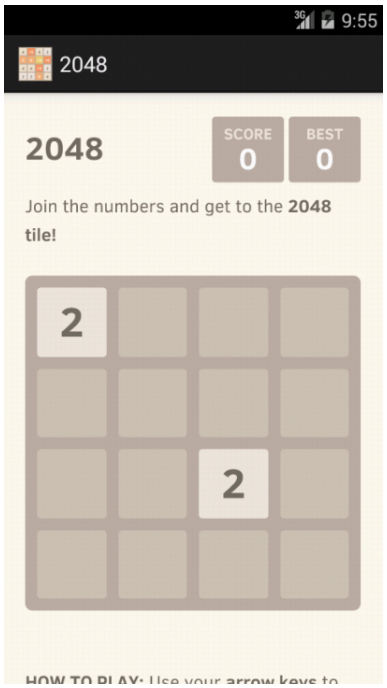

Figure 4. Example of the game 2048 user interface. Image taken from the Google Play page of the application.

repetition software. At first user engagement can be seen as a subjective concept, thus difficult to measure. One first approach might be to perform a qualitative analysis of the application. To do so, the type of required information would be related to the perception of the user about the application. Such information might have the form of opinions, comments, and suggestions.

Collecting qualitative data can be expensive since it might require to interview the participants after using the application. Alternatively, surveys or questionnaires could be utilised to collect that type information. However, the biggest problem is that this kind of information is difficult to analyse. Moreover, the information might be biased by the mood of the participants and other uncontrolled factors when collecting it. Under such conditions, it might be highly difficult to replicate the evaluation.

Some methods have been developed to cope with the difficulty of measuring user engagement. One of them is the development of a framework to measure user engagement (O'Brien & Toms, 2010) by assessing parameters like perceived usability, aesthetics and felt involvement. Such assessment is performed through a series of surveys. Even though this option tries to minimize the problems related to collecting and analysing qualitative data, its main problem is that it is a general purpose framework that needs a considerable amount of information. Thus, it might be difficult to adapt it when the number of participants is small.

Due to the difficulty pertaining to gathering and analysing qualitative data, the evaluation for the current project will be done in terms of quantitative data. This will ease the collection of data and their analysis. Moreover, it will permit to replicate the evaluation or perform a new one with modified conditions in order to make comparisons. The details of the evaluation method will be described in the following subsections.

3.1. Participants

There will be two groups of participants. The first one will be the control group. The participants in this group will use a version of AnkiDroid with no gamification characteristics. The data obtained from this group will be used as a benchmark to measure the variation in user engagement. The second group will be the experimental group. The participants in this group will use the gamified version of AnkiDroid.

The results obtained from the experimental group will be compared against the benchmark from the control group. The expectation is that there will be differences between both results. The analysis of those differences will allow to draw conclusions about the effectiveness of gamification for spaced repetition software.

The number of total participants will depend on the availability of subjects to participate in the evaluation of the final application. The expectation is to have between seven and ten people in every group. In any case, the number of members in both groups will be the same so the amount of collected data will be comparable.

AnkiDroid is a general purpose implementation of spaced repetition. Thus, it can be used for any person that have access to an Android device. In fact, the published application in Google Play falls in the content rating *Everyone*, so there is no restriction in the age of its users. Such situation means that the constraints in the selection of the participants are minimum.

Even though the minimum constraints in the selection of participants, it is important to maintain a level of homogeneity among them. Thus, the complexity in the analysis of the results from the control and experimental will be further reduced. A potential issue with this condition is the level of generalizability of the results. It might be difficult to extrapolate the results to other groups of the general population.

Finally, the participants will use the final application at any time during their daily activities. It follows that the application will be used in an uncontrolled environment. There are several potential issues for such condition. However, the most relevant one is that the user might forget to use the application, therefore not enough data will be generated. To diminish such situation, the application will sent daily notifications to remember the users to use it. Evidently, the frequency of such notifications has to be moderated to avoid the users perceive them as intrusive.

3.2. Quantitative data

The approach to measure the effects of gamification applied to a spaced repetition software will be done through quan-

titative data. There are two parameters to measure. The first one is the user engagement, which can be measured by collecting and analysing different types of data from the application including time of use, frequency of use, and commonly used features. The second parameter is the effectiveness of spaced repetition when it is implemented with gamification. This parameter require to test how much knowledge has been retained by the participants after using the application.

3.2.1. User engagement

There are three types of data to be collected from the usage of the application:

1. **Time in session (TIS):** It is the time the users spend every occasion they use the application. It is measured in minutes per session.
2. **Frequency of sessions (FOS):** It is the time between sessions. It is measured in hours between two consecutive sessions.
3. **Frequency of spaced repetition features (FOSR):** It is the number of times spaced repetition features are used in every session. It is measured in number of times per session.

3.2.2. Spaced repetition effectiveness

The data to measure the spaced repetition effectiveness will be collected during the usage of the application and in a final test to be taken by the participants. Such information will take the form of following scores:

1. **Score in session (SIS):** The score obtained in the quiz taken during a session of usage. It is measured as the number of correctly answered questions divided by the total number of questions.
2. **Final score (FS):** The score obtained in the final test. It is measured as the number of correctly answered questions divided by the total number of questions.

3.3. Hypotheses

The main objectives of the project are to increase the user engagement of AnkiDroid, and measure the effectiveness of spaced repetition when implemented with gamification. Thus, the following hypotheses need to be confirmed or denied:

H1: The participants in the experimental group will spend more time using the application than the participants in the control group.

H2: The participants in the experimental group will obtain higher scores than the participants in the control group.

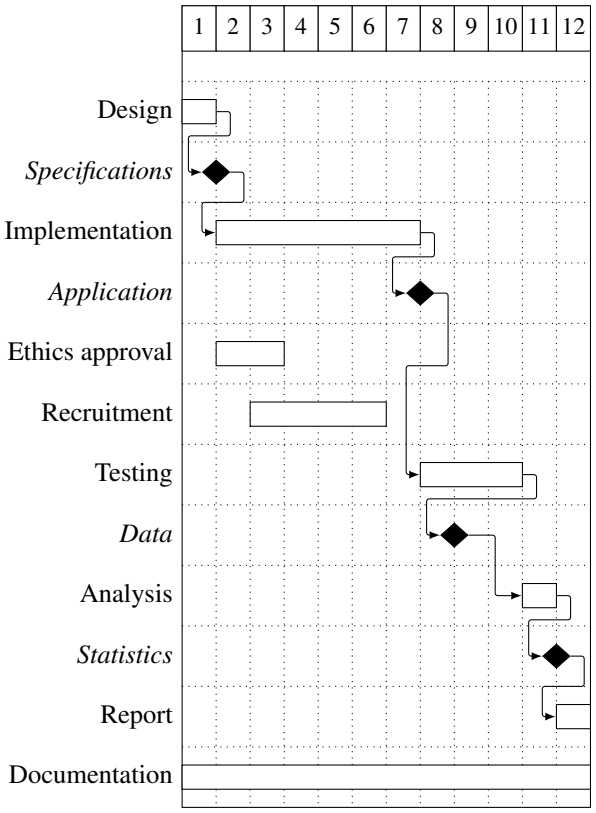

Figure 5. Gantt chart of execution of project tasks over a period of 12 weeks. Some tasks deliver milestones needed to start the following task.

3.4. Analysis

The analysis of the collected information will be aimed to confirm or deny the hypotheses to some level of confidence. It will require to make proper statistical analysis in the data from both groups. Initially, the analysis will be made using student's t-test and ANOVA, although other methods might be required.

4. Tasks and Schedule

The execution of the project is defined as a series of tasks or stages with specific goals. Each task has a number of resources, a duration, and outcomes that might be required in a following stage as seen in Figure 5. The next subsections describe the details of each task.

4.1. Design of the final application

An overview of the general aspects of the design has been described previously (2). However, specific details about use case, application features, user interface components, and user interactions are still required. The objective of such details is to clearly establish the scope and context of the final application. Those details will be described in a specifications document.

Based on the specifications document, a first prototype of the user interface will be created on paper. Then, the

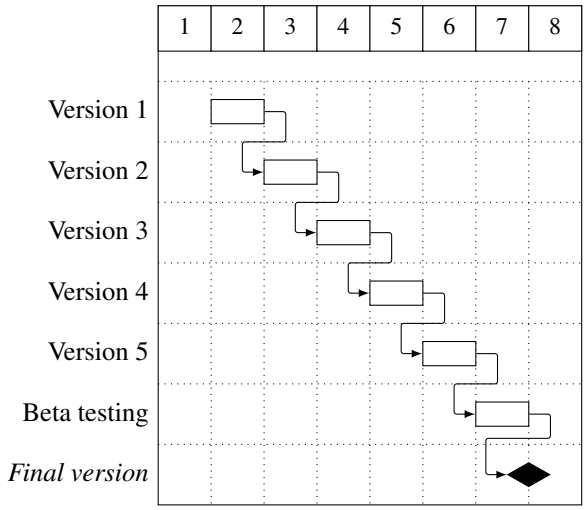

Figure 6. Gantt chart of implementation stage. An iterative process where each phase adds new features and fixes bugs from previous versions.

prototype will be evaluated using a cognitive walkthrough approach (Helander, 1997) to detect potential flaws in the design for first time users. All of the participants in the testing stage will be considered first time users since the application will be new. The result of this analysis will be used to improve the design of the user interface.

The execution of this task is critical since its results will be the foundations for the implementation and analysis stages. The specifications document will contain not only the details of the application from the perspective of the user, but it will also contain the types of information to be collected to perform the analysis, and how they will be collected and stored. The outcome of this stage will be a specifications document containing details about the features of the application along with sketches of the user interface. The duration of this task will be one week.

4.2. Implementation of the design

This stage corresponds to implementing the application based on the output from the previous stage. This stage is divided into iterative phases as seen in Figure 6. At the end of each phase, an updated version of the application will be delivered. Each new version will contain additional features and will fix bugs from the previous ones.

Each one of the iterative phases is expected to last one week. The total number of phases will be five. An additional iteration lasting one week will be needed to perform beta tests. The implementation stage will end once the beta tests have been executed and passed, and the final version of the application has been released.

4.3. Ethics approval for research

Since testing the application will require the participation of humans, it will be necessary to obtain the corresponding approval from The School of Informatics. The objective of this stage is to make sure that the activities during the testing stage will follow the guidelines from the School Ethics Code and ethics regulations at the College and University, therefore, no harm will be caused to the participants. This stage will require to fill and submit a Level 1 Ethics Procedure form.

Additionally, the details of the project, and the consent form to be signed by the participants will be provided. Both elements will be created out of the specifications document from the design stage. The purpose of the consent form is to make sure the participants understand the implications of the study. This task will need to be done early during the development of the project, ideally, after the specifications document is ready. The expectation is that the study will be approved within two weeks.

4.4. Recruitment of participants

Once the final version of the application has been released, it will have to be tested by the participants in the control and experimental groups. At this point the participants will have been recruited already. Moreover, they will have to be aware about the purpose of the testing stage and their role. Such scenario poses a contact with potential participants as early as possible.

Ideally, potential participants will be contacted during the first steps of the implementation stage when a basic functional version of the application is already released. The purpose is to be able to provide screen-shots of the interface of the application. Thus, the potential candidates will have a rough idea of the look and feeling of the application. The expectation is that this strategy will encourage potential candidates to participate in the study.

Finally, the actual participants will have to be contacted before the release of the final version of the application. It follows that the participants will be fully aware of their role, the objective of the testing, the duration of the testing period, and the information that will be collected. To do so, they will have to sign the consent form previously approved by The School of Informatics. This process is expected to last four weeks. The outcome will be the consent forms signed by the participants.

4.5. Testing of the final application

The participants will test the final version of the application. As mentioned, there will be two groups: control and experimental. This stage will be conducted as a blinded experiment, thus, the participants will not be informed to which group they belong to. The purpose is to avoid potential biases due to the users knowing what the expected outcomes are. After signing the consent form, they will be given the resources to install the application.

The testing stage will last between two and three weeks. The actual duration will depend on the amount of data collected after the second week of study. During this time, the participants are expected to use the application in a

daily basis. Daily notifications will be sent remembering to use the application. The data from the application will be collected remotely as long as the participants are connected to a free wifi service to avoid charges in their mobile service due to the use of the application.

The testing stage will require the participants to take a final quiz. Such quiz will be taken within the application. The participants will be notified remotely about the final quiz. After completing that quiz, the participants will be informed that testing period has ended and they will be free to uninstall the application. Once the testing period has ended, the participants will receive a monetary compensation. The outcome of this stage is a set of data collected while the participants used the application.

4.6. Analysis of data

The data collected during the testing stage will be used to perform analysis related to the user engagement and spaced repetition effectiveness. Since, the data will be gathered remotely, it will be possible to start the analysis before the end of the testing stage. The expectation is that in the middle of the testing stage there will be enough data that can be used to find trends or other characterizations in the usage of the application.

The analysis process can be speed up by performing automated routines to clean and pre-process the data on a daily basis. Once all the data have been collected, a simple processing will be required. The outcome of this stage will be a series of statistical parameters that characterize the usage of the application in the control and experimental group. This information will be used to confirm or deny the hypotheses. This stage is expected to last one week.

4.7. User engagement and spaced repetition effectiveness report

The outcome from the analysis of data will be the foundations to draw conclusions about the effects of gamifying a spaced repetition software. The report will be aimed to interpret the results from the analysis data and identify possible causes for such results. Moreover, information about the limitations of the project will be provided. Either the hypotheses have been confirmed or denied, the report will provide the guidelines for further work in the same field or related ones. This stage is expected to last one week.

4.8. Documentation

This task will run through the entire period of the project. It starts along with the design stage, and it lasts until the end of the project. The objective is to describe the details of the execution of each task. It will provide a clear way to understand how the project was developed including changes, problems, and solutions that might have found throughout the process.

5. Risk Management

As seen in Table 1, each stage has a defined duration and outcomes. Some stages depend on the completion of previous ones as seen in Figure 5. It follows that a delay or a problem in one stage could potentially affect the execution of the following ones, hence, the flow of the entire project.

To diminish the affectation of the project due to issues in every stage, it is important to identify and classify potential risks. Moreover, strategies to mitigate the effects of such risks are needed. Table 2 provides a summary of the potential risks and mitigation strategies. The following subsections provide more details about risks in every stage and how to minimize their influence in the flow of the project.

5.1. Design of the final application

The risk in this stage is that the design misses some important aspects that will affect the interaction of the users with the application. The severity of this risk is high since it would affect all subsequent stages (implementation, testing, analysis and report.) As previously described, the alternative to identify flaws and improve the design is to perform an evaluation of a first prototype.

A single evaluation of the first prototype of the application can miss some important aspects as well. Minimizing the effects of such situation will require the application to be evaluated by at least two human computer interaction experts. The recommendations from those experts will be utilised to create the final design of the interface.

5.2. Implementation of the design

The iterative phases of this stage are meant to implement new features and fix bugs. Therefore, there are two types of risks. The first one is that a high number of features could delay each iteration. The severity of this risk is medium since some features can be avoided. Ideally, the features will be categorized by their importance and relevance to the application. Then, the most important ones will have to be implemented during the first iterations. If at the end of the fifth iteration some secondary features are not implemented yet, they will have to be discarded.

The second potential risk is the creation of new bugs when fixing previous ones or implementing new features. Similarly, the severity of this risks is medium since features causing bugs can be discarded as long as their relevance is small. The effects of this risk can be minimized by the creation of test cases that will have to be ran when a new feature is added or a previous bug is fixed. Alternatively, a test driven development (TDD) strategy can be adopted. It is important to note, that the creation of test cases can be time consuming, therefore, it is necessary to keep a balance between them and the overall implementation process.

Stage	Duration (weeks)	Outcomes
Design	1	Specifications document
Implementation	6	Final version of the application
Ethics approval	2	Approval from The School of Informatics
Participants recruitment	4	Signed consents
Testing	2-3	Data
Analysis	1	Statistics
Report	1	Report
Documentation	12	Project documentation

Table 1. Duration and outcomes of every stage of the project.

Stage	Risks	Severity	Mitigation
Design	Miss important aspects	High	Double evaluation of prototype
Implementation	Too many features	Medium	Categorization and prioritization
	Bugs	Medium	Test cases & TDD
Ethics approval	Not approved	Low	Guidelines & feedback
Participants recruitment	A few participants	High	Early contact
Testing	Low usage	High	Notifications
	Abandonments	Medium	Extra participants
Analysis	Incorrect analysis	Low	Advice from experts
Report	Unproper conclusions	Low	Advice from experts
Documentation	Delay in other stages	Low	Daily updates & weekly revision

Table 2. Summary of potential risks in the project.

5.3. Ethics approval for research

The major potential risk in this stage is that the study has not been approved in the first submission. The severity of this risk is low since the task is executed along with the implementation stage and its duration is small. This risk is minimized by following the guidelines provided by the School Ethics Code and ethics regulations at the College and University. If the study is not approved in the first submission, it is expected that feedback and recommendations will be provided to improve the next submission.

5.4. Recruitment of participants

Participants are the most important component of the testing stage. No participants mean that the testing stage is not possible to execute. The potential risk in this stage is that the number of participants is not enough to collect sufficient data for the analysis stage. The severity of this risk is high since not enough data will affect the confirmation or denial of the hypotheses. The effects of this risk are minimized by contacting the potential participants early in the development of the project.

5.5. Testing of the final application

Once the testing stage has started, the usage of the application will depend on the participants only since they will use it at some points during their daily routines. The main risk is that they will forget about the application, therefore, not data will be collected. The severity of this risk is high since the lack of enough data will affect the posterior analysis. The first strategy to cope with this risk is the implementation of daily notifications to remember the participants to use the application.

The second risk is that the participants abandon the study. Evidently, participants will not be forced to continue in the study. To reduce the affectation of this risks, an extra number of participants will be needed, as long as the budget allows it. Moreover, the monetary compensation will be given at the end of the study. Finally, the consent form must establish that abandoning the study means no monetary compensation will be granted.

5.6. Analysis of data

This stage the data will require to make a statistical analysis. The risk here is that the proposed analysis might not provide the results to confirm or deny the hypotheses. The severity of this risk is low since at this point there will be more alternatives to analyse the data. One of those alternatives will be looking for advice from experienced people in the field. Moreover, the vast amount of resources for data analysis will help to have a better understanding of the collected information.

5.7. User engagement and spaced repetition effectiveness report

This is the last stage of the project. At this point everything else have been already done. This means that the effects of the risks at this stage will be low. One risk might be the drawing of not proper conclusions. This effect can be minimized by getting advice from experts in the fields related to the project.

5.8. Documentation

Documenting the stages poses a risk to the flow of the whole project. Writing the documentation could be time consuming, therefore, it would affect the execution of the tasks. The strategy to minimize those affectations is to update the documentation on a daily basis when possible. At the end of every week a revision will be done. If is not possible to add all the details related to the activities done during the week, then the task will be skipped until the next week. During the testing stage most of the activities will be related to the participants which will provide some spare time to update the documentation and fill any gap from the previous weeks.

References

Boticki, Ivica, Baksa, Jelena, Seow, Peter, and Looi, Chee-Kit. Usage of a mobile social learning platform with virtual badges in a primary school. *Computers & Education*, 86:120–136, 2015.

Cheong, Christopher, Cheong, France, and Filippou, Justin. Quick quiz: A gamified approach for enhancing learning. In *PACIS*, pp. 206, 2013.

Damien, Elmes. Anki, 2018. URL http://apps.ankiweb.net/.

DiSalvo, Briana B Morrison Betsy. Khan academy gamifies computer science, 2014.

Gené, Oriol Borrás, Núñez, Margarita Martínez, and Blanco, Ángel Fidalgo. Gamification in mooc: challenges, opportunities and proposals for advancing mooc model. In *Proceedings of the Second International Conference on Technological Ecosystems for Enhancing Multiculturality*, pp. 215–220. ACM, 2014.

Graafland, Maurits, Schraagen, Jan M, and Schijven, Marlies P. Systematic review of serious games for medical education and surgical skills training. *British journal of surgery*, 99(10):1322–1330, 2012.

Hamari, Juho and Koivisto, Jonna. âĂIJworking out for likesâĂİ: An empirical study on social influence in exercise gamification. *Computers in Human Behavior*, 50: 333–347, 2015.

Hamari, Juho, Shernoff, David J, Rowe, Elizabeth, Coller, Brianno, Asbell-Clarke, Jodi, and Edwards, Teon. Challenging games help students learn: An empirical study on engagement, flow and immersion in game-based learning. *Computers in Human Behavior*, 54:170–179, 2016.

Helander, Martin G. *Handbook of human-computer interaction*, pp. 717–732. Elsevier, second edition, 1997.

Hintzman, Douglas L. Theoretical implications of the spacing effect. 1974.

Johnson, Matthew W, Okimoto, Tomoko, and Barnes, Tiffany. Leveraging game design to promote effective user behavior of intelligent tutoring systems. In *International Conference on Intelligent Tutoring Systems*, pp. 597–599. Springer, 2012.

Marcus, Burnette. Flipquiz, 2018. URL https://flipquiz.me.

Michael, David R and Chen, Sandra L. *Serious games: Games that educate, train, and inform*. Muska & Lipman/Premier-Trade, 2005.

O'Brien, Heather L and Toms, Elaine G. The development and evaluation of a survey to measure user engagement. *Journal of the Association for Information Science and Technology*, 61(1):50–69, 2010.

Salomon, Gavriel. The differential investment of mental effort in learning from different sources. *Educational Psychologist*, 18(1):42–50, 1983.

Salomon, Gavriel. Television is" easy" and print is" tough": The differential investment of mental effort in learning as a function of perceptions and attributions. *Journal of educational psychology*, 76(4):647, 1984.

Slish, Donald, Nash, Micael, and Premo, Joshua. Gamification increases scores of underperforming students in cell biology. In *EdMedia: World Conference on Educational Media and Technology*, pp. 870–876. Association for the Advancement of Computing in Education (AACE), 2015.

Su, C-H and Cheng, C-H. A mobile gamification learning system for improving the learning motivation and achievements. *Journal of Computer Assisted Learning*, 31(3):268–286, 2015.

Uberspot. 2048, 2017. URL https://play.google.com/store/apps/details?id=com.uberspot.a2048.

von Ahn Luis, Hacker Severin. Duolingo, 2018. URL https://www.duolingo.com/.

Zichermann, Gabe and Cunningham, Christopher. *Gamification by design: Implementing game mechanics in web and mobile apps*. " O'Reilly Media, Inc.", 2011.