Game-based learning vs gamification for language learning

Final report

15.2.2021

Authors

Shanshan Hou <shahou@kth.se>

Ondřej Brém <ondrejb@kth.se>

Abstract

Utilizing game elements and components in the learning process are popular in digital education. This paper focuses on the comparison of learning experience between game-based learning and gamification. We chose two language learning applications, Duolingo and Influent, which each represent gamified learning and game-based learning software respectively. We divided our users into two groups, participating in our experiments, surveys and interviews. Based on the data we gathered from participants, they had more fun and playfulness and also better memory retention in case of Influent.

Table of Contents

List of Acronyms

HCID - Human computer interaction design

HCD - Human centered design

UX - User experience

UI - User interface

MMO games - Massively multiplayer online games

1. Introduction

When it comes to learning languages there are many methods that students can use. Outside of traditional textbooks and courses of various forms, one can use different digital tools. Especially with the rise of smartphone usage the possibility to have an app for language learning became very attractive. It gives the learner the possibility to learn at their own pace and whenever they have time, regardless of the schedule of any classes.

In this project we explore those learning tools, in particular we are looking at gamification as a tool to help with the learning process and game-based learning. By conducting an experiment we will show how each of the methods influences the experience of the learner during their studies. We will show comparison of aforementioned approaches through an experiment with two groups of students.

Both methods are designed to improve engagement and motivation of students. In this project we are exploring a possible way to compare gamification and full game-based learning. We will select two example applications for this experiment to show the process on them as well as show the basic comparison of the methods.

2. Research Questions

For our project we stated following research questions:

- **RQ1:** How does gamification and game-based learning compare in terms of learners' experience.
- **RQ2:** Is there a different result in terms of efficiency of memorizing words between the two methods?

Our hypothesis is that game-based learning should perform better as it doesn't focus solely on external motivation in the form of rewards but gives the learner the joy from the play itself.

3. Literature Study

With the goal of comparing the learning experience of gamified learning with game based learning we first need to take a look at what is the difference between these two and look at the HCD methods that can be used for user experience research. Additionally we need to take a look at existing apps and platforms that would best suit as representative examples for our experiment.

3.1. Gamification

Gamification takes game elements (such as points, badges, leaderboards, competition, achievements) and applies them to a non-game setting (EdTechReview, 2013). For instance, a lot of social media platforms use a progress bar when the users need to fill in a profile which is actually based on the sense of progress in a game. Research shows it increases the profile completion by 30% (Karl, 2019). Gamification is mainly used for enhancing motivation, engagement and attention in education (Dichev & Dicheva, 2017) by turning routine, mundane tasks into refreshing, motivating experiences (EdTechReview, 2013). Other

examples of gamification include language learning apps like Duolingo[1] and Language Drops[2] or platforms like Khan Academy[3] that use elements of gamification to help with motivation and progress tracking. Another example, Kahoot![4] can be likely considered as an example from both categories. It's mainly using points and other gamification methods but at the same time the setting in which it is often used is not exactly plain teaching but more of a game play in itself. Moving on to the Game based learning we can see that the border between the two is fairly blurry.

3.2. Game Based Learning

According to definition by (EdTechReview, 2013), game based learning is a type of game play that has defined learning outcomes. Generally, game based learning is designed to balance subject matter with gameplay and the ability of the player to retain and apply said subject matter to the real world. Compared with gamification, points can also appear in game based learning, but this method takes more importance in creating playful experience (Plass & Homer, 2020). As such game based learning might not feel and look like a learning experience at first glance. Both physical games, board games and video games can belong to this field. For instance, in the famous game Angry Bird[5], users need to utilize different angels and powers to destroy pigs, in Crayon Physics[6] participants need to draw shapes in suitable positions to gather points, both these two games are related to physics. Even for MMO games, such as League of Legends[7] and World of Warcraft[8], players need to know how to cooperate well, make strategies carefully, communicate effectively and operate accurately to win the game. Compared with what mentioned above, the board game Trivial Pursuit[9] is a more game based learning process, players review old knowledge and gather new knowledge while playing.

4. Methods

To achieve our goals of comparing gamified learning and game based learning of languages, we need to select two example applications to be compared. For the gamification example it is pretty easy to find suitable platforms such as history, language, math or trash sorting games as this method is commonly used in digital learning applications. While for the game based learning example, as the real game-playing feeling is very important in our experiment, we utilize more criteria to value them: game experience, test environment, device requirement and time cost (Figure 1). Finally, we chose Duolingo and Influent[10], each of them represents one of the approaches, gamification and game based learning respectively. For each of the applications we selected five participants for the study. We also selected French as a language of choice because it was available in both of the selected

[1] https://www.duolingo.com/learn
[2] https://languagedrops.com/
[3] https://www.khanacademy.org/
[4] https://kahoot.com/
[5] https://www.angrybirds.com/
[6] http://crayonphysics.com/
[7] https://na.leagueoflegends.com/en-us/
[8] https://worldofwarcraft.com/en-us/
[9] https://sv.wikipedia.org/wiki/Trivial_Pursuit
[10] https://store.steampowered.com/app/274980/Influent/

tools. Another benefit of using French over other options was it used Latin alphabet which everyone speaking english, our base language, already knows.

Category	Name	Game experience	Test environment	Device requirement	Time cost
History	Civilization V	good	computer	high	high
	Assassin's Creed	good	computer	high	high
Language	KLOO Race	normal	in-person	normal	low
	Influent	good	computer	normal	low
Common knowledge	Trival Pursuit	normal	xbox, ps4	normal	low

Figure 1

The research was conducted in a few steps. We started with an experiment where participants used each respective app for a short time, followed with questionnaire and debrief interview to get more open-ended questions answered. Due to COVID situation all the experiments and interviews were conducted remotely over video calls.

4.1. Participant Selection

When selecting participants for our experiment we needed to decide who is our focus group. As (Mortensen, 2020) explains: In user testing of a product or service, future or current users of said service are a natural choice because they are the ones for whom the design is being done. With this research we are testing a general concept and the goal is not to improve either of those applications. Because of that the focus group is selected more arbitrarily to serve as an example for this research. In our plan we established requirements for our participants that we needed to meet for our experiment:

- Participants should be involved in a learning process
- For comparability between the groups we will focus on student age participants
- For practical reasons the participants need to speak English
- Participants need to be familiar with using computers and smartphones so that additional issues outside of the tested experience don't influence the results

We assembled two groups of participants to test the two learning methods. The groups were formed in a way to be as similar as possible to avoid skewing of the results by participant selection. By doing so we were able to compare the results of both of the methods between each other.

For the practicality of this experiment we recruited the participants from the students of KTH and Aalto University which in a way limits the generality of our results but as well allowed us to have somewhat similar participants in the experiment. This approach also partially helped with having students with similar backgrounds in terms of study experience, which hopefully helped with avoiding some skewing of the results.

4.2. Experiment

The experiment itself is mainly going to be based on user research methods, mainly moderated and unmoderated testing (Moran, 2019). In our case we are not testing any

particular UI for the sake of improving it but rather an approach to learning. In the role of facilitator we will need to take this into consideration and adjust questions and tasks for the participants accordingly.

The experiment was divided into two parts corresponding to the respective learning app. As said we recruited two groups of five students and each performed the learning in their assigned app. In both cases we focused on the very early stage of language learning. Our participants were thus selected to be novices in French and did use the apps to learn the very first French lesson.

4.2.1. Influent

Influent is a 3D game aimed at teaching multitudes of languages by using a virtual environment where the words are placed with the items. The players walk around the environment, a small flat in the first French lesson and collect words that they want to learn. After collecting enough words a practice is offered. This can be either finding the items on time or flying around with a model plane and shooting corresponding items based on the word given by the game.

In our experiment we let the participants play the game for about 15 minutes, during which they collected and practiced 20 words. As mentioned the experiment was conducted remotely over Zoom. Participants were left to navigate the game and select the words as they wished. From the facilitators perspective only observations and occasional support, when technical issues arose, was performed. The players were encouraged to ask if something is unclear as the time for the experiment was quite limited.

Most of the participants used an instance of the game installed on their own computer and provided the list of collected words after the experiment. One instance was conducted via screen share in zoom where the game ran on our computer.

4.2.2. Duolingo

Duolingo is a language learning app that uses gamification principles like rewards and levels to teach a multitude of languages. It has both a web app and mobile app. It has the teaching divided into topics where each topic has 5 levels and each level has several exercises. Players can continue to the next level only after they finished the previous level and new topics get unlocked only when previous ones are finished at least at level one.

In our experiment we let our participants go through the first level of topic “Basics 1” to make it quite comparable in complexity with the beginning of Influent game-play. To track progress of each participant we created 5 new accounts that we distributed to them. The participants will use the Duolingo website or application in their own devices.

4.3. Debrief

After the experiments were conducted, participants were given a questionnaire to fill, which was designed to quantify their experience. On top of the questionnaire, a structured interview was conducted to gain more qualitative insight into how they perceived the experience of learning via respective methods.

In the questionnaire we tried to understand how well each respective method supported the learning process of the participant as well as their motivation. This was conducted by a list of statements and giving the participants an agreement scale to answer on. The whole questionnaire is available in appendix.

During the short structured interview we tried to better understand the context of our participants and get a bit of background to help us with understanding thor answers from the questionnaire. Additionally it was an opportunity to learn interesting observations and specific reasons for why they answered in a certain way in the questionnaire.

4.4. Memory Test

To evaluate the result of the learning process, we prepared a vocabulary test for our participants. During the experiment the words that each of the participants were learning and practicing, were recorded. Next personalized test was created and distributed to each respective participant to test what they remember from that lesson. Comparison between results from each app was conducted to see if there are any visible differences.

5. Results Analysis

After conducting the experiment collected data was analyzed from two perspectives. One focusing on user experience of the learning and the other looking at the memory effects of each method. Our quantitative survey is simple and concise, which mainly works for the rating of user experience and memory retention. As we use Google Form to conduct our survey, we also use it for analysing results.

5.1. Experience

5.1.1. Influent

Based on the experiment with Influent we gained few observations, which were to some degree confirmed by the answers in the questionnaire. Starting with the first question from the questionnaire asking about how participants perceived the experience we can see a slight tendency towards the play side, but not conclusive.

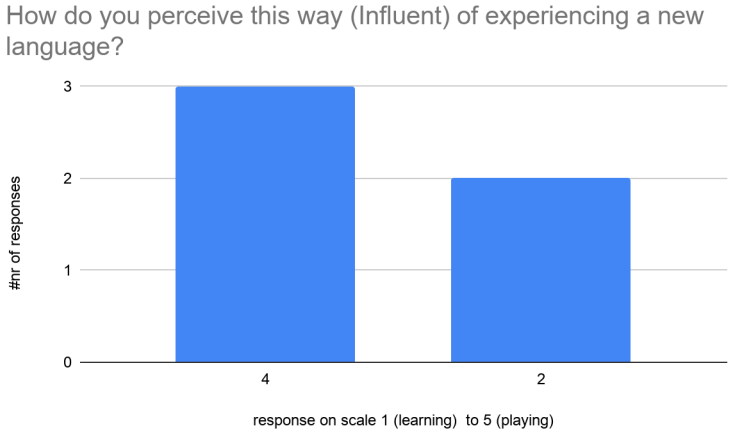

Figure 2

But we can add to that observation that the participants seemed to enjoy themselves during the experiment and mentioned it was fun during the interview. Further in the questionnaire 4 out of 5 answered that it was fun to play the game. Same result followed for a question if they enjoyed the experience.

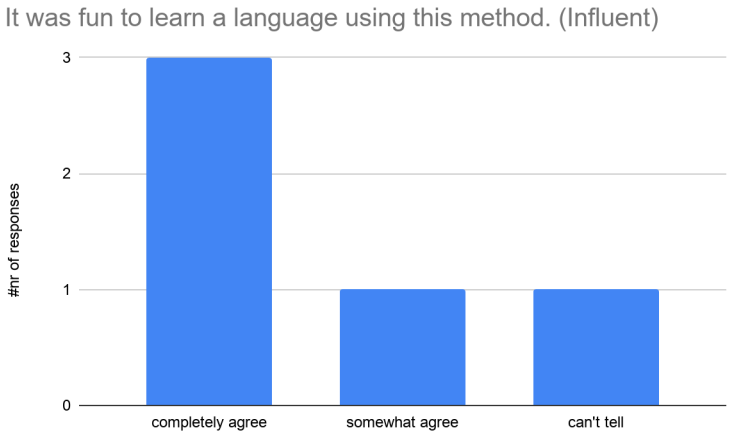

Figure 3

Interestingly 4 out of 5 participants that learned using Influent said they would use the method again for language learning. Last participant likely wouldn't use it again. Surprisingly all participants agreed to some extent that this approach to learning a new language supports their learning process. With those results in mind it's rather unsurprising that most of the participants would prefer this method over traditional textbook based language study at least to some extent.

During the interviews participants often mentioned too complicated controls for the game and some poor UI elements that made understanding the interactions more difficult. Another thing that was hindering the learning process was pretty poor feedback on errors as mentioned by the participants. One of the participants even said that they think they didn't yet get to actual learning during the experiment.

5.1.2. Duolingo

From quantitative result:

When talking about the perception of the experience, our four participants gave four different values from 1 to 4. Our participants all showed high confidence in the learning experience.

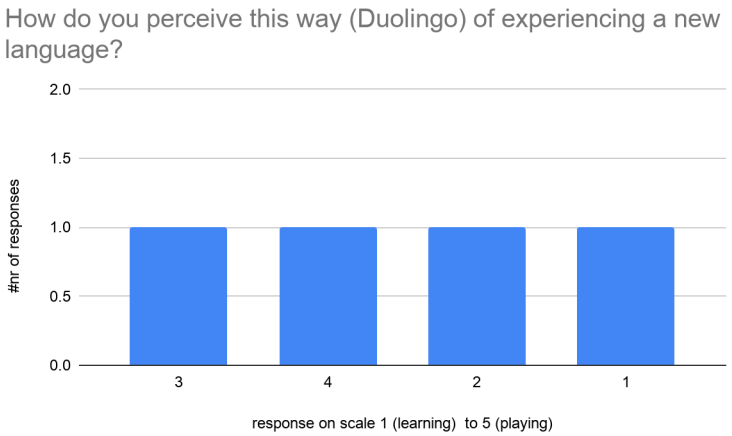

Figure 4

No one felt the game elements disturbed their learning experience, also the feedback to correct errors is easy to accept. The most convergent opinions appear on whether it is easy to learn from mistakes from Duolingo, half of our participants agreed and half of them disagreed. Most participants felt fun and enjoyed the learning process.

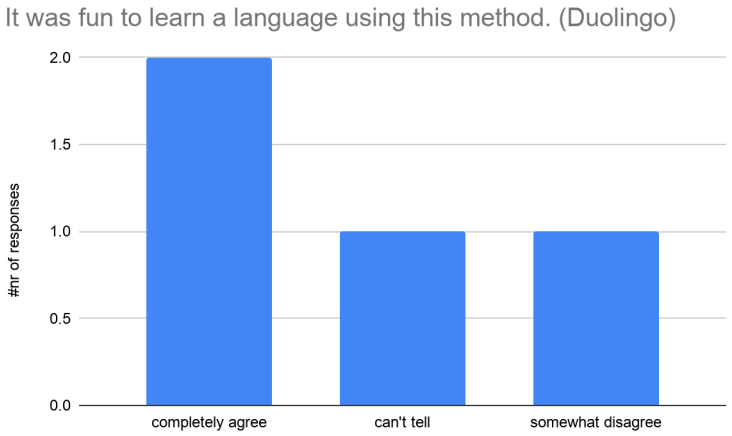

Figure 5

From qualitative research:

We assumed that the Duolingo test would be easier to make while it was even harder to recruit participants than Influent. People felt it should be boring when they heard the application, which made us think will students' attitude and behavior changed by their understanding of the tool.

Our participant felt the level setting is reasonable and the difficulty gaps fit well to a beginneR. The visual style and animation make the learning process interesting. The limited hearts made our user treat the game more seriously and even added a sense of competition. One participant mentioned she felt the learning process is a little bit slow. Onr participants felt it was more for beginners but not for professional learning.

5.1.3. Comparison

When comparing Influent and Duolingo in terms of learning experience perceived by participants of our experiment we ended up focusing on three questions.

1. Did the participants feel more like learning or playing?
2. Would they use the respective tool again?
3. Was it fun to learn using the respective app?

Following chart shows that Influent was only slightly more perceived as a play compared to Duolingo. In both cases the responses were evenly distributed around the midpoint of options. With this little amount of datapoint it is thus not possible to draw any conclusions.

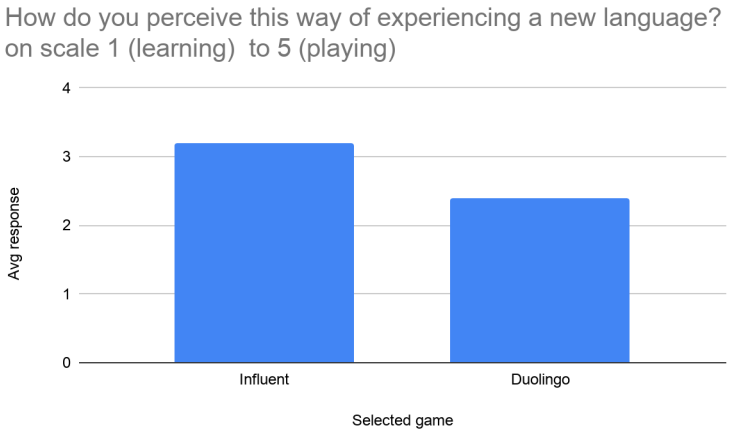

Figure 6

Next if we look at the fun factor Influent again scored better, not by high margin though.

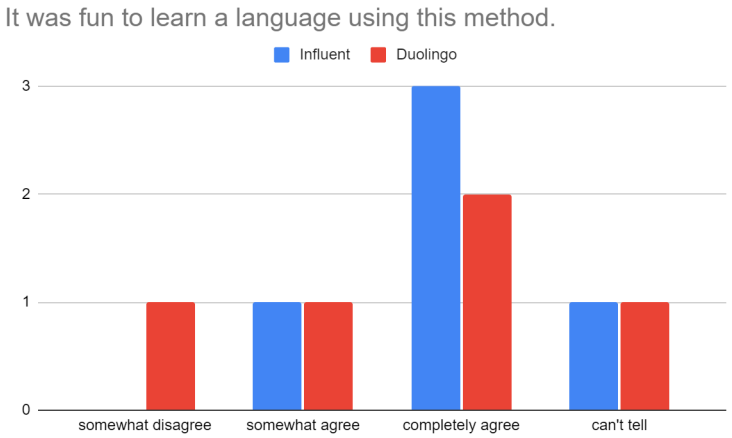

Figure 7

Lastly we looked at willingness to use each respective app again for learning a new language. Here again Influent got better reception. From corresponding interviews this can be attributed to more engagement provided by the gameplay and the participants having more fun trying out this less traditional approach to language learning.

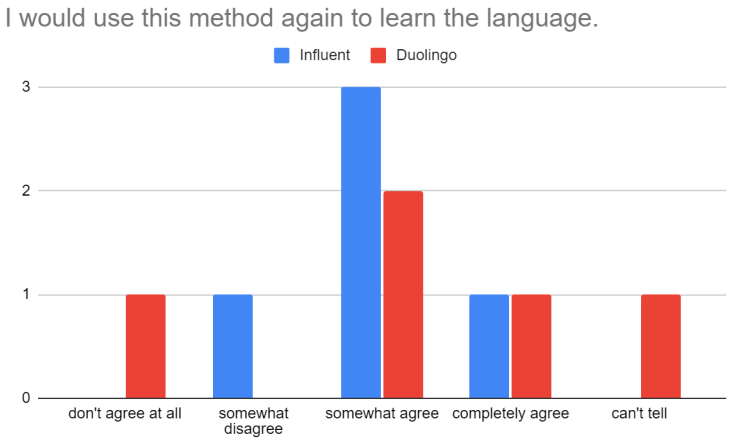

Figure 8

Overall the experience was perceived similar in both apps with slightly more fun and playfulness in case of Influent. Considering the scale of the experiment this is not to be taken as conclusive evidence (one or two changed responses could drastically change the results) but more as an indication where to look further and what to expect. Meanwhile, the learning experience can also be influenced by how long they experienced those applications. For example, users' interest might get reduced after they have played the game for a long term which might lead to the changes in 'learning-playing' feeling.

5.2. Learning

In the learning part of the experiment we asked the participants to answer a quick memory test a week after they played with the respective app. In the test they were given a set of words that were part of their learning exercise to see how many of them they correctly remember. They were given 10 words in total, 5 and 5 in each translation direction. Each correct response was given a point and those were then counted to see the performance.

- Influent points: 3,4,4,7,10
- Duolingo points: 1,1,7

As evidenced by points gained in each of the games Influent shows better results. It is however important to note the scale of the study wasn't large enough to drive reliable results. Additionally as in Influent the players have the possibility to select the words to learn there is the issue of them selecting words that are either familiar to them or they already know from some previous learning or game.

On the other hand Duolingo is missing a couple data points as we didn't get the needed responses in time or at all to include them in the study. The late responses were dismissed as well because the time difference between the participants would likely affect the memory effects.

6. Discussion

First and foremost we want to make clear that results from this research are only indicative and their reliability is limited due to low number of participants and use of only selected

learning apps. Further the setup of the experiment couldn't simulate properly the real conditions and context in which learners would normally use these tools to learn a new language. This last point is mainly due to recruiting and the limited resources to find proper participants.

Starting with the recruiting issue, our participants were willing to take part in the experiment but didn't necessarily have any prior interest in learning French or any new language for that matter. With that being the case, their motivation coming into the experiment wasn't aligned with the aims of the activity. To fix this, it would be best to find participants that are already interested in learning a new language and letting them learn a language of their choice. This way we would eliminate the biases originating from the participants inherent disinterest in language learning.

Next issue that we realized and was pointed out by at least one of the participants was the inherent short duration of the activity due to the experiment. In case of Influent this leads to the participants spending much of the time just figuring out the mechanics and controls of the game which distracts them from the learning experience itself. Additionally the insights into the experience and motivation of the participants were only short term and could differ vastly after longer usage. This was again pointed out by one of the participants when they said it would likely get boring quickly to play Influent unless the environment changes or more maps are available.

Specifically in case of Influent we ran into some issues with implementation of the game and UX of the controls which made the experience worse for the participants but didn't necessarily have any relation to game-based learning in general. Thus for better results it would be better to do this study on a larger scale using more different applications and extrapolate from there the results that are related to the game-based method of learning and separate those issues that are implementation specific.

Lastly we want to point out issues in the memory part of the experiment. As evidenced by missing responses from Duolingo participants we ran into issues getting answers for our memory tests. Some of them we didn't get at all and some of them took longer to get back. This disparity in time could cause differences in the results that aren't visible. The main reasons for this were difficulties in getting the participants again after set time.

In the end we would like to suggest a few things to focus on for future work and experimentation. First and foremost a larger number of participants for the study would provide higher statistical significance for any of the questions answered in the study. Secondly if the learning outcomes are to be part of the study, longer duration of the experiment is needed to properly see any effects and to avoid biases and "cheating" of the participants during the learning process.

To further improve the data gathering, collaboration with the app providers would be desirable for direct collection of data about the learning outcomes and experience of the players. Additionally to get a generalized comparison between gamification and game based learning more different applications should be used to eliminate the direct effect of UI and implementation of any specific app. To further generalize the comparison from language learning to other areas of learning it should be considered to experiment with other fields as well.

Taking our experience and applying it to such expanded study would provide more significant results and more generalized answers to our research questions.

7. References

Dichev, C., & Dicheva, D. (2017). Gamifying education: what is known, what is believed and what remains uncertain: a critical review. *Dichev and Dicheva International Journal of Educational Technology in Higher Education*, *14*(1), 9. https://link-springer-com.libproxy.aalto.fi/content/pdf/10.1186/s41239-017-0042-5.pdf

EdTechReview. (2013, April 23). *What is GBL (Game-Based Learning)?* EdTechReview. Retrieved October 8, 2020, from https://edtechreview.in/dictionary/298-what-is-game-based-learning

Interaction Design Foundation. (n.d.). *User Research*. Interaction design foundation. Retrieved 10 9, 2020, from https://www.interaction-design.org/literature/topics/user-research

Karl, J. (2019, 8 12). *Games Vs Gamification: What's the Difference?* YouTube. Retrieved 10 8, 2020, from https://www.youtube.com/watch?v=kHn0Maj8ygs

Moran, K. (2019). *Usability Testing 101*. Nielsen Norman Group. Retrieved 10 6, 2020, from https://www.nngroup.com/articles/usability-testing-101/

Moran, K., & Pernice, K. (2020, 4 26). *Remote Moderated Usability Tests: How to Do Them*. Nielsen Norman group. Retrieved 10 9, 2020, from https://www.nngroup.com/articles/moderated-remote-usability-test/

Mortensen, D. H. (2020). *The Basics of Recruiting Participants for User Research*. Interaction design foundation. Retrieved 10 9, 2020, from https://www.interaction-design.org/literature/article/the-basics-of-recruiting-users-for-usability-testing

Plass, J. L., & Homer, B. D. (2020). Theoretical Foundations of Game-Based Learning. In *Handbook of Game-Based Learning*. MIT Press. https://books.google.se/books?id=_2fKDwAAQBAJ&lpg=PR5&lr&pg=PR5#v=onepage&q&f=false

Robinson, K. (2006, February). *Do schools kill creativity?* TED. Retrieved 10 6, 2020, from https://www.ted.com/talks/sir_ken_robinson_do_schools_kill_creativity?language=en

8. Appendix

8.1. Questionnaire

8.1.1. Live version

https://forms.gle/LR14ruMJMMJgJJRM6

8.1.2. Questions

- How do you perceive this way of experiencing a new language?
    - Learning to Playing scale of 5
- How confident do you feel about your learning during this session?
    - Not at all to I know all scale of 5
- Rate each of the following statements in terms of how much do you agree with them (don't agree at all, somewhat disagree, somewhat agree, completely agree, can't tell)
    - It is easy to learn from mistakes with this method.
    - I would recommend this method to my friends.
    - I would prefer learning languages this way compared to using textbooks.
    - Learning using this method met my expectations.
    - It was easy for me to learn a language using this method.
    - I would use this method again to learn the language.
    - I enjoyed the experience.
    - The game elements didn’t disturb my learning
    - It was fun to learn a language using this method.
    - The feedback to correct errors is easy to accept.
    - This method supports my learning process.

8.2. Interview questions

Interview questions were formulated loosely based on the following list that was prepared before the experiment. The topic of the questions was kept but sometimes more explanation was needed to get to the answers.

- How would you describe (evaluate) the experience you just had?
- Do you currently learn any new language?
- What was the best approach for learning languages for you in the past?
- Is there any specific technique that helped you with learning languages in the past?
- Consider the method you just experienced for language learning. Can you mention some positives and negatives in it for you personally?
- General comments depending on where the interview went.