Master's thesis

NTNU
Norwegian University of Science and Technology
Faculty of Information Technology and Electrical Engineering
Department of Computer Science

Sigurd Greiff
Jørgen Hollum

Supporting Social interaction, Repetition & Question Sequence in a QA Platform

Master's thesis in Computer Science
Supervisor: Trond Aalberg
January 2023

NTNU
Norwegian University of
Science and Technology

Sigurd Greiff
Jørgen Hollum

Supporting Social interaction, Repetition & Question Sequence in a QA Platform

Master's thesis in Computer Science
Supervisor: Trond Aalberg
January 2023

Norwegian University of Science and Technology
Faculty of Information Technology and Electrical Engineering
Department of Computer Science

NTNU
Norwegian University of
Science and Technology

NTNU
Kunnskap for en bedre verden

Department of Computer Science

Master Thesis

Supporting Social Interaction, Repetition & Question Sequence in a QA Platform

Authors:
Sigurd Greiff
Jørgen Hollum

Supervisor:
Trond Aalberg

Spring, 2023

Abstract

In recent years, computer-based learning has become a frequent method of acquiring and repeating knowledge. While these digital solutions offer enhanced engagement and motivation for learners, they often lack the social interaction inherent in face-to-face discussions and collaborations. Closed questions commonly used in computer-based question-answer (QA) systems limit the opportunity for applied self-explanation and deeper understanding.

The primary objective of this Master's thesis is to explore the integration of vocal social interaction with computer-based QA learning, emphasizing knowledge retention and utilizing question sequencing. To achieve this, we investigate the effects and ways of including social interaction, examine the influence of question order and difficulty, and explore methods for supporting recall and memory retention. Employing an iterative design process, we conducted user testing, incorporating observations and qualitative interviews for evaluation.

Our findings demonstrate that the inclusion of open-ended questions, coupled with a democratic voting system to classify vocally provided answers, and allowing an administrator to guide the pace of the learning experience, facilitates discussions and enhances social interaction. We discovered that the order in which questions are presented sparks user engagement, influenced by factors such as interleaved or blocked practice and the player's current performance. Additionally, we leveraged question difficulty to establish personalized learning paths that challenge users appropriately, fostering motivation and creating a flow state. To support recall and memory retention, we provided statistical recommendations and repetition questions, encouraging users to strive for better performance and engage with the learning material repeatedly.

By incorporating open-ended questions and a democratic voting system, educators can enhance social interaction and foster deeper learning experiences. Furthermore, by considering question order and difficulty, personalized learning paths can be created to optimize individual engagement and motivation. Tailored repetition and analysis, within computer-based learning systems, can effectively enhance learning effectiveness.

Sammendrag

De siste årene har databasert læring blitt en hyppig metode for å ta til seg og repetere kunnskap. Mens disse digitale løsningene tilbyr økt engasjement og motivasjon for elever, mangler de ofte den sosiale interaksjonen som kommer fra diskusjoner gjennom ansikt-til-ansikt og samarbeid. Lukkede spørsmål som vanligvis brukes i digitale spørsmål-svar-systemer begrenser muligheten for anvendt selvforklaring og dypere forståelse.

Hovedmålet med denne masteroppgaven er å utforske hvordan man kan integrere vokal sosial interaksjon med digital spørsmål-svar-læring, som har fokus på å bevare kunnskap og sekvenseringen av spørsmål. For å oppnå dette undersøker vi effektene og måtene for å inkludere sosial interaksjon, undersøker påvirkningen av spørsmålsrekkefølge og vanskelighetsgrad, og utforsker metoder for å støtte hukommelse og kunnskapsbevaring. For å finne ut av dette tok vi i bruk en iterativ designprosess med brukertesting som inkluderte observasjoner og kvalitative intervjuer for evaluering.

Våre funn viser at å inkludere åpne spørsmål, kombinert med et demokratisk avstemningssystem for å klassifisere vokalt gitte svar, og muligheten for en administrator til å styre læringstempoet, skaper diskusjoner og forbedrer sosial interaksjon. Vi oppdaget at rekkefølgen spørsmålene blir presentert i vekket brukerengasjement. Dette var påvirket av faktorer som vekslende eller blokkert øving og spillerens nåværende prestasjoner. I tillegg utnyttet vi spørsmåls vanskelighetsgrad for å tilpasse seg brukerens kunnskap som videre utfordrer brukerne på en optimal måte, fremmer motivasjon og skaper en tilstand av høy konsentrasjon og glede. Hukommelse og kunnskapsbevaring ble tilrettelagt gjennom statistiske anbefalinger og repetisjonsspørsmål gitt. Dette oppfordret brukere til å prestere bedre og engasjere seg med læringsmaterialet gjentatte ganger.

Ved å inkludere åpne spørsmål og et demokratisk avstemningssystem kan pedagoger forbedre sosial interaksjon og læringsopplevelser. Videre kan spørsmål tilpasses ved å vurdere rekkefølge og vanskelighetsgrad for å optimalisere individuell engasjement og motivasjon. Skreddersydde repitisjonsspørsmål og analyser, i digitale læringssystemer, kan forbedre læringsresultater.

Acknowledgments

This thesis concludes our Master of Science degree, in Computer Science, at the Norwegian University of Science and Technology in Trondheim, Norway. The research has been supervised by Trond Aalbarg and conducted during the spring semester of 2023. We are thankful for the guidance, constructive feedback, and support you provided, steering us in the right direction along the way.

Lastly, we would like to give a thank you to the individuals that participated in the user test. The feedback and insights contributed significantly to the research and realization of this master thesis.

Contents

List of Figures

List of Tables

1 Introduction

This thesis investigates how social interaction, order of questions, and repetition can be leveraged in a computer-based question-answer (QA) learning system. In recent years, computer-based learning has become a frequent method of acquiring and repeating knowledge. As these systems tend to prioritize computer interaction over social interactions learners might miss opportunities for prompting self-explanation in discussions. In this research, we aim to find techniques that address this challenge and support memory retention.

The following introductory section aims to provide background information and context for the Master's thesis. The research questions to be addressed in this research are presented, and the chosen research method is explained and justified. Furthermore, an outline of the report is provided, giving an overview of the thesis's structure.

1.1 Motivation

In the current century, technology has had a significant impact on the learning environment. It has rapidly evolved to be an important tool to optimize learning in classrooms through research and modernization. Today, the traditional pen and paper have been replaced with smart- tablets, phones, and laptops due to their accessibility. Digital software, tailored for educational environments, has emerged as a means to deliver high-quality learning experiences and tries to address challenges related to student motivation and engagement through playful elements. As technology continues to reshape the learning landscape, it is essential to consider the role of social interaction in optimizing educational experiences.

Repeating and learning different subjects in a group setting is a well-established practice across a variety of educational contexts. These types of settings rely on social interaction to acquire knowledge and to help motivate learners. In physical settings, learners have the advantage of being able to engage in direct communication with their peers and effortlessly convey their emotions through facial expressions and body language prompting engagement. Unfortunately, these expressions are often overlooked in computer-based learning, as they interact with computers instead of humans.

With the increasing availability of educational digital resources, traditional educational practices have undergone a significant shift. One popular educational method that has been part of this shift, in QA systems, is quizzes. Quizzes have long been recognized as an effective tool for assessing knowledge, that lets users create engaging and interactive learning experiences. However, the digital adaption of quizzes contains closed questions and sometimes multiple-choice answers in order to simplify the process of validating answers. This lacks the opportunity of using more open questions and create social interactions. This leads to a lack of space for constructing and elaborating own responses to questions and the possibility of learning from co-learners.

In addition to the limitations of existing QA systems, many of these solutions overlook the potential benefits of utilizing both question difficulty and order to tailor question sequences, thereby optimizing motivation and engagement. Existing research acknowledges that the difficulty and order of questions impact learning outcomes. Questions that are either too easy or too difficult can negatively impact learner motivation and engagement, ultimately affecting overall learning outcomes. By exploiting question order, appropriate questions can be provided for each participant. When multiple subjects are incorporated into a QA session, the order of questions can offer a deeper understanding of the relationship between them [5].

While existing research emphasizes the importance of repetition for knowledge retention [8], the process of repeating material can often be demotivating. Therefore, it is important to explore different mechanisms that engage users in repetition and motivate learners to repeat material. By building upon existing techniques that effectively engage users and encourage material repetition, our research aims to create a more engaging approach that inspires and motivates learners throughout the repetition process.

As previously stated, repeating and learning material in physical group settings is a well-established practice. These settings naturally possess the social dynamics that are desired to incorporate into a digital solution. In essence, the intention is to incorporate these social dynamics to develop a QA platform that highlights the importance of repetition, provokes social interaction, and adapts question sequence while focusing on making the process fun and engaging.

1.2 Research Questions

As discussed in the motivation chapter, our primary objective of this Master's thesis is to combine computer-based QA learning and vocal social interaction that emphasizes knowledge retention and utilizes question sequence. Additionally, the study aims to expand current research on learning technology and seeks to contribute to the collection of options available for computer-based learning. To guide the research we have set the following three research questions:

RQ1: How can social interaction and computer-based learning be combined for an interactive QA system?
The inclusion of social interaction within classroom settings allows learners to actively engage with one another and benefit from shared knowledge. Moreover, computer-based learning has demonstrated its efficacy in delivering valuable educational content. These two approaches hold the potential for a powerful combination that can enhance learning experiences. To assess its potential, we will examine how to combine them for an interactive learning QA system and evaluate its effects.

RQ2: How can question difficulty and order of questions be utilized?
The interpretation of a question's difficulty level may vary depending on an individual's knowledge level. This research question is raised in order to explore how this variability can be leveraged, in a QA system, to provide questions that are most suitable for an individual's proficiency level and arrange them in a sequence that

maximizes their learning outcomes.

RQ3: How can support for recall and memory retention be implemented?
It is known that the retention of newly acquired knowledge decreases with time. In order to address this challenge, we pose this research question to investigate techniques that facilitate memory recall and retention, ensuring that the learned information remains accessible and retained for an extended period. By supporting memory-enhancing strategies, individuals can prevent the decay of knowledge retention and maximize the long-term benefits of their learning efforts.

1.3 Research Method

To address our research questions and accomplish our main objective, we will create a platform centered around these inquiries. To achieve this, we will follow the process illustrated in Figure 1.

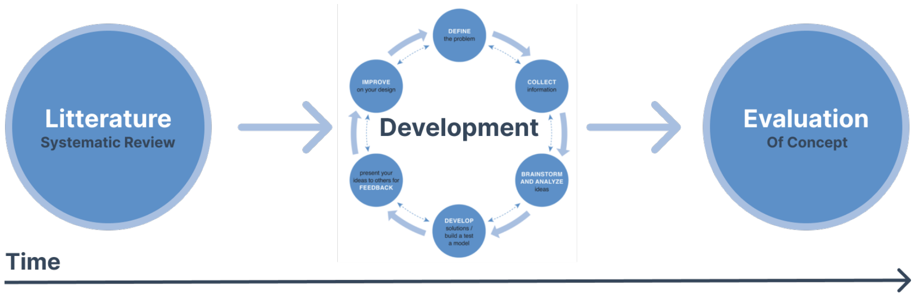


Figure 1: An overview of the process utilized

The initial step involves exploring existing literature and solutions related to the subject matter. This exploration enables us to generate ideas and techniques that can work as a foundation for development.

The development will be conducted through an iterative design process that involves identifying issues and requirements in our implementation, as illustrated in Figure 2. This offers a structured approach to research and promotes continuous improvements. Through a user-centered design, we can explore new innovative solutions that may enhance the relevancy and impact of the research. Furthermore, documenting the design process enhances the transparency and reproducibility of the research.

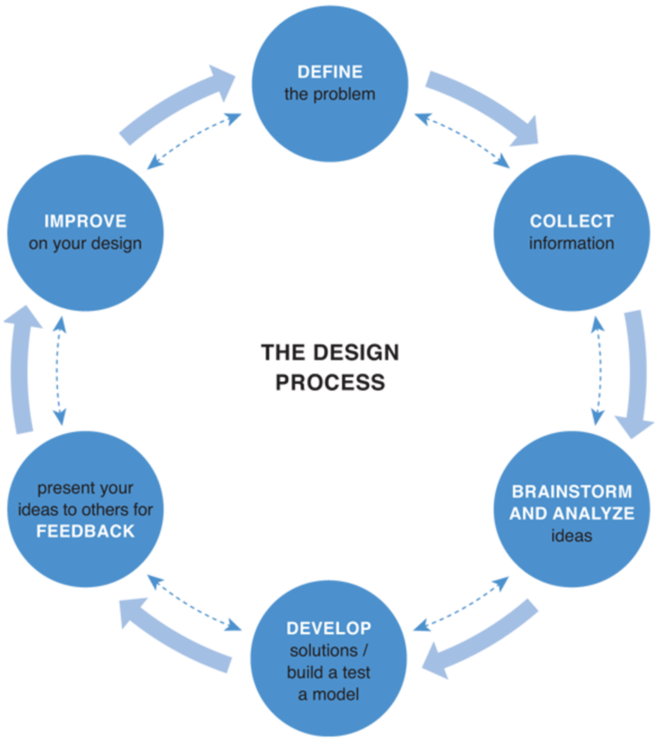


Figure 2: Illustration of the design process followed in the thesis [3].

Upon identifying and defining problems, we will brainstorm and analyze the collected information before creating our first prototype. The first prototype will be a low-fidelity paper prototype in order to test central concepts [25]. The problems and needs discovered will be incorporated when designing a high-fidelity prototype using computer software. By creating a high-fidelity prototype we can test more specified areas of the concept and improve the design before starting on creating the platform. Once the platform is finished, we can proceed to the final step to conduct the evaluation.

In order to evaluate the effectiveness of the techniques and ideas implemented in the proposed solution, we will employ a qualitative approach. This approach involves presenting test cases to users, followed by semi-structured interviews to facilitate in-depth discussions and gather valuable insights. This provides us with observations, user experiences, perceptions, and interactions through the perspective of learners. The interviews can further illuminate thoughts, preferences, and suggestions in a more personalized manner. Overall we can assess the effectiveness of our mechanisms, make informed improvements, and inform responses to the research questions.

1.4 Report Outline

The following sections of this thesis are structured to provide a comprehensive understanding of the research process.

Section 2: Background provides an overview of the theory, concepts, techniques, and existing solutions relevant to the research questions. These topics are essential for comprehending the ideas and techniques employed to tackle the challenges associated with the research questions.

Section 3: Design presents and discusses proposed ideas and techniques that could be incorporated into the solution, building upon the theoretical foundation established in the previous section.

Section 4: Implementation outlines the prototyping process, offering insights into the steps involved, including brainstorming, low- and high-fidelity prototypes, and the development of the proposed solution.

Section 5: Evaluation aims to present observations and insights derived from qualitative interviews conducted during user testing. The findings are aligned with the research questions and examine the impact of the implemented techniques and features in addressing these questions.

Section 6: Conclusion & Further Work presents the final contribution of the master thesis, providing a conclusion to the research questions. This section also acknowledges the research's limitations and offers suggestions for future work.

Minor elements in the background and design have been reused from previous project work.

2 Background

Effective learning techniques are essential for enhancing knowledge and understanding. This thesis aims to optimize learning outcomes by combining computer-based QA learning with vocal social interaction, prioritizing knowledge retention, and strategic question sequencing. Therefore, it is important to look at the theoretical background related to these goals and identify gaps in existing solutions.

To establish a strong foundation for the proposed solution, it is crucial to delve into the theoretical background related to learning, repetition, engagement, and motivation. This section will examine the significance of engaging and effective learning techniques and underscore the potential of gamification in educational contexts. By delving into these theories and concepts, valuable insights can be gained to enhance learning outcomes and support the development of our solution.

2.1 Repetition of Learning Material

Repetition is crucial for retaining information and increasing knowledge retention. However, repetition techniques can sometimes become monotonous, and it is essential to find engaging and useful ways to make the process enjoyable. To develop a solution that is effective, it is necessary to understand the key elements of learning and repetition.

2.1.1 Retrieval Practice & Self Explanation

Retrieval practice involves recalling to-be-remembered information from memory. This can typically be recalling material presented by a teacher in a course or taking a quiz from the last lecture. In the context of a student taking a quiz, without any aids, he or she must recall the topic from memory in order to get the right answer. Washington University [16] experimented on eighth-grade students in middle school to see whether recalling information through quizzes would improve their final grade in science. Over the course of a year and a half, quizzes were spaced out and as a result, it produced significant learning benefits with between 13% and 25% gains in performance on the total unit examination. By manipulating the placement of extra quizzes before and after the lecture, and as a review prior to the unit exam they saw that this produced the greatest increases in exam performance. These increases were only slightly improved when the items had appeared on previous quizzes.

Students who explain examples to themselves tend to learn better, make more accurate self-assessments of their understanding, and use analogies more economically while solving problems [30]. Researchers Kiran Bisra et al. [4] found that having learners generate an explanation is often more effective than presenting it. The most powerful application of self-explanation is when the learner revises their explanation due to new information, highlights, gaps, or errors.

2.1.2 Forgetting Curve

In order to implement support for recall and compensate for memory retention, it is important to get a general understanding of how the brain stores knowledge and how it decays. This can be a useful tool for both developing an interactive QA system and for providing insights and feedback to motivate and engage users.

In 1885 Hermann Ebbinghaus performed a self-study to clarify how memory decays over time and how repetition affects it [8]. The study has been successfully replicated in modern times [10] and indicates that people tend to continually halve their memory of newly learned knowledge in a matter of days unless they repeat the learned material. Most of the forgetting occurs within the first minutes and hours of learning [24]. After a day or two, we typically forget around 3/4 of what we have learned. At this point the retained knowledge drops at a slower pace, meaning that we will retain only a few details but we will be able to remember them for several more days. The original study and modern replication also show that by repeating the learned materials we can increase retention, and also decrease the time spent relearning the material. An estimate of memory retention, with n repetitions, can be described mathematically through the following formula [35]:

$$R(n) = exp^{\frac{-k*t}{S[n-1]}}$$

- R(n) = Retrievability at the n-th repetition
- k = Decay constant
- t = Time in intervals
- S[n-1] = Stability after the (n-1)th repetition

2.2 Social Interactions and Learning

Learning is a significant part of any process of acquiring knowledge. Learning can occur through various methods, including formal education, hands-on experiences, observation, and interaction with others. With the advent of technology, learning has become more accessible and personalized, offering a multitude of digital resources, online courses, and interactive platforms. The ability to learn and adapt is essential in today's rapidly evolving world, where continuous learning is necessary to keep up with changing demands and advancements in various fields.

People often turn to others for learning and the social interaction plays an important role in learning. In a social setting, learning can occur by comparing ourselves to peers, or observing others to develop a better understanding of the subject [20]. In a study of reciprocal teaching [21], they found that students may compare their understanding to what they observe in another person and think more deeply about who is right. This implies that observing a peer can trigger learning and reflections [20].

2.2.1 Spaced & Massed Learning

Spaced learning is when a learner reviews a given material over time with breaks between the sessions. The learner repeats a material reminding him or her of prior occurrences. This presentation of previous occurrences enhances the memory process. Several cognitive and educational psychology studies have demonstrated that spacing out repeated encounters with the material over time produces superior long-term learning, compared with repetitions that are massed together [12]. Incorporating tests into the spaced practice is proven to amplify the benefits.

Contrary to spaced learning, with session breakups, massed learning does not implement any breaks, making it a continuous session. Massed repetition eliminates the retrieval process due to the fact that the material was just presented, hence there is no need to retrieve it from memory [31].

2.2.2 Blocked- & Interleaved Practice

Blocked practice groups the same kind of material together during practice, as illustrated in Figure 3, e.g. AAABBBBCCCCDDDD. With blocked practice, on the contrary, as all the problems require the same strategy, the student needs only focus on executing a given strategy repeatedly [12]. Carvalho, Paulo F. and Goldstone, Robert L. [5] find that blocked practice improves learning of low similarity categories by increasing within-category comparison.

Interleaved practice is achieved when different subjects are intermixed during practice e.g ABCDABCDABCD where two intervening materials (e.g., C and D) come before each recurrence of a given item (e.g., A). This is further illustrated in Figure 3.

Sean H. K. Kang [12] finds that interleaving with more, rather than fewer, intervening items further increases the learning outcome. Switching between types of problems also improves the learners association between problem type and its strategy, which promotes successful problem-solving. This is further supported by Carvalho, Paulo F. and Goldstone, Robert L. [5] who finds that interleaving improves discriminative contrast between different categories, while studying each concept in separate blocks emphasizes the similarities within each category. Interleaved studies have been shown to improve the learning of high-similarity categories by increasing between-category comparison.

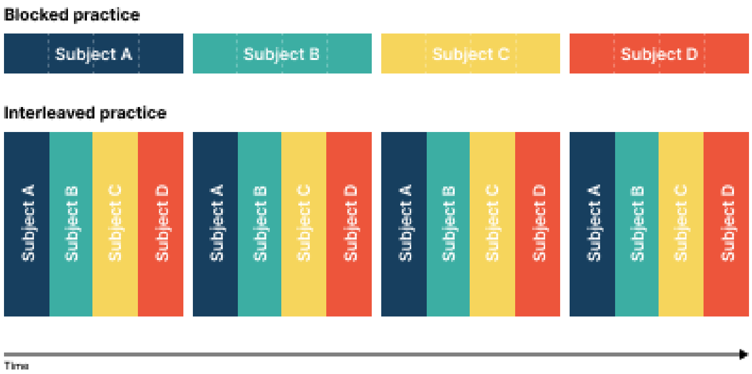


Figure 3: Illustration of interleaved- and blocked practice in a spaced context

2.3 Flow Theory

Any kind of learning activity requires engagement in order to be successful. Including social interaction in computer-based learning requires users to be engaged and enjoy themselves, otherwise, the social dynamic will be minimal. Therefore, finding out how people get engaged in an activity and how to replicate it is important in order to engage participants.

Csikszentmihalyi's flow theory [6] was developed when attempting to understand how people feel when they most enjoy themselves. The study showed that enjoyment did not necessarily come from relaxing or living without stress, but rather during activities in which their attention was fully absorbed. They explain the flow state with expressions such as "everything clicks" [6].

The flow theory associated with learning suggests that there are three states: boredom state, flow state, and frustration state [34]. When the learner performs tasks, the learner's flow state can be preserved if the task's difficulty matches the learners developing skills. If the task difficulty fails to increase with the learner, the result will be the boredom state. This is where the learner is not interested in the task and disengages from the activity because it is too easy. On the contrary, if the tasks increase in difficulty faster than the learner, the result will be the frustration state. The tasks within this state may still be of interest, but the tasks become so challenging that they lose motivation to continue. Therefore, the goal of the activity is to fall between these states, called the flow state. This is where the task difficulty increases with the learner but does not succeed the learner. This way the learner's interest and challenge are maintained over time [6] [27].

2.4 Assessment System

An assessment system is a collection of assessments where each represents a tool for tracking what and how well students have learned [33]. Washington, DC: The National Academies Press [33] propose that such a system should be composed of both assessments designed to support classroom teaching/learning and self-monitoring purposes. This provides learners with a basis for awareness, reflection, and new insights.

2.5 Quiz

A quiz can be used to test a person's understanding or knowledge of a subject. It is often used as a form of entertainment, but can also be educational. In a quiz, participants are asked a series of closed questions and must answer correctly in order to score points.

Ayanna K. Thomas et. al [28] conducted a study to see whether frequent quizzes in a college course pay off. The class engaged in three different interval practices, namely quizzes, quizzes with feedback, and study-based. The study demonstrated that when the in-class practice was retrieval-based, students performed better across different test formats and on new but related concepts, than when the practice was study-based. As a result, interval quizzing is highly beneficial to test with multiple choice and short answer questions as well as post-collaboration facilitation.

2.5.1 Question Difficulty

There are several methods to maintain a challenging environment [2]. (1) The difficulty can be fixed, but the study's researchers found that this does not lead to skill development or engagement. (2) The difficulty gradually increases with the learner's advancements, which encourages engagement and increases learning performance. (3) The difficulty is adapted based on the learner's performance, in which the learner can be engaged with the environment created by their skill level.

2.5.2 Question Validity

A study by the University of South Australia on multiple-choice questions using a facility index and discrimination efficiency index claims that it is possible to identify questions that potentially need to be modified [22].

Facility Index indicates how easy a question was for students by calculating the percentage of learners who answered a question correctly [26]. If less than 5% of learners answer the question correctly, the question is either extremely difficult or might be formulated in a way that makes it difficult to understand. If the question gets a facility index lower than 30 % it should be reviewed [22].

Moodle is the world's most widely used learning platform with more than 213 million users [17] as of 2020. The platform has established a facility index [18], presented in Table 1. The statistics are designed for summative tests, where students have just one attempt for each question.

Discrimination Efficiency Index is that those whom in general performs well

Facility index	Interpretation
5 or less	Extremely difficult or something wrong with the question.
6-10	Very difficult.
11-20	Difficult.
21-34	Moderately difficult.
35-65	About right for the average student.
66-80	Fairly easy.
81-89	Easy.
90-94	Very easy.
95-100	Extremely easy.

Table 1: Moodle's facility index [18]

on quizzes, has a tendency to answer other questions correctly as well. A low discrimination efficiency index indicates that there is something wrong with the question. If a learner that in general answers incorrectly suddenly answers correct on question X, and learners that usually performs well answers incorrectly on question X, will lead to a low discrimination efficiency index. Questions with a discrimination efficiency index below 20 should be reviewed [22].

2.6 Motivational & Educational Game Design Principles

Motivation can be achieved extrinsically- and intrinsically. Extrinsic motivation is connected to rewards from external sources such as grades, money, and avoidance of punishment. Intrinsically motivation, on the contrary, looks at the contributions to the person's genuine enjoyment and satisfaction when performing an activity. This can for example be achieved through challenge, curiosity, and cooperation.

Gamification is a way to utilize extrinsically- and intrinsic motivation through the use of game elements and design in non-related game contexts. The goal of gamification is to increase motivation, and user activity, and encourage particular behaviors. Elements present in games that have been mainly used in educational systems design include scores, leaderboards, achievements/badges, and levels [7].

Laine, Teemu H. and Lindberg, Renny S. N. [14] conducted a systematic literature review on game design principles (DPs) and game motivators with a specific focus on games for education. In total, there were found 56 motivators categorized into 14 classes (based on their similarities), and 54 DPs categorized into 13 classes. The 13 classes, in the context of DPs, are provided in Figure 4, and their corresponding motivators are highlighted in italic font.

Challenge (*challenge*, *control*, *curiosity*, *feedback*, *goals*)

- DP 1: Provide challenges at various adjustable difficulty levels
- DP 2: Favor simple challenges over complex challenges
- DP 3: Provide enough time to solve challenges
- DP 4: Raise curiosity by interesting and/or unpredictable challenges
- DP 5: Allow challenges to be repeated

Control (*control*, *customization*, *feedback*, *novelty*, *real world relation*, *sensory curiosity*)

- DP 6: Use input modalities that are suitable for target players and contexts
- DP 7: Use sensitive and accurate controls
- DP 8: Use consistent controls
- DP 9: Use familiar, safe, and comfortable controls
- DP 10: Favor simple interaction
- DP 11: Use adjustable and personalized controls
- DP 12: Freedom of choice and control in gameplay

Creativity (*competence*, *control*, *curiosity*, *customization*, *fantasy*, *recognition*)

- DP 13: Allow players to create game content
- DP 14: Provide creative ways to solve challenges
- DP 15: Allow players to extend the game system

Exploration (*control*, *curiosity*, *discovery*, *navigation*, *physical movement*, *real world relation*)

- DP 16: Freedom of exploration and experimentation
- DP 17: Provide multiple paths/options and dynamic ordering of events
- DP 18: Provide an interactive game map of sufficient detail

Fairness (*challenge*, *competence*, *moral sense*, *rules*)

- DP 19: Implement mechanisms to prevent or detect cheating
- DP 20: Ensure similar opportunities of success regardless of experience

Feedback (*control*, *competence*, *emotions*, *feedback*, *goals*, *novelty*, *rules*, *sensory/cognitive curiosity*, *usefulness*)

- DP 21: Provide instructions and/or tutorials
- DP 22: Provide immediate, positive, and useful feedback
- DP 23: Provide clear feedback via different channels
- DP 24: Provide access to performance data

Goals (*challenge*, *competence*, *fantasy*, *feedback*, *goals*, *rules*, *status*, *usefulness*)

- DP 25: Create clear, meaningful, and achievable goals
- DP 26: Provide an epic meaning
- DP 27: Create progressive goals that build on each other

Learning (*cognitive curiosity*, *competence*, *feedback*, *immersion*, *real world relation*, *usefulness*)

- DP 28: Provide relevant and pedagogically grounded learning content and activities
- DP 29: Provide cognitive challenges
- DP 30: Embed assessment tools
- DP 31: Choose an appropriate transparency level for learning content
- DP 32: Provide time to reflect amid intense game moments

Profile and ownership (*competence*, *control*, *feedback*, *goals*, *recognition*, *status*)

- DP 33: Use a profile/avatar that the player can own and can relate to
- DP 34: Provide the status of the game process and next available actions
- DP 35: Offer past, present, and future perspectives

Relevance and relatedness (*control*, *real world relation*)

- DP 36: Relate gameplay to real-world contexts
- DP 37: Relate to familiar activities
- DP 38: Relate to past experiences

Resources and economy (*cognitive curiosity*, *competition*, *control*, *recognition*, *rules*, *social interaction*, *tension*)

- DP 39: Develop a virtual economy to allow trade and utilization of resources
- DP 40: Enable collection of virtual goods
- DP 41: Make some resources scarce
- DP 42: Introduce a possibility of losing resources

Social play (*control*, *collaboration*, *competence*, *competition*, *recognition*, *relatedness*, social interaction, *status*)

- DP 43: Provide means for social communication and interaction
- DP 44: Provide opportunities for competition
- DP 45: Provide opportunities for collaboration
- DP 46: Provide means for expressing status and recognition among players
- DP 47: Allow players to form and manage groups
- DP 48: Allow players to distinguish themselves in their group

Storytelling and fantasy (*control*, *customization*, *emotions*, *escapism*, *fantasy*, *immersion*, *real world relevance*)

- DP 49: Create a meaningful story that the player can relate to
- DP 50: Provide a fantasy context
- DP 51: Offer role-play experiences
- DP 52: Create possibilities for players to interact with the story
- DP 53: Enable user-created stories
- DP 54: Create thought-provoking scenarios

Figure 4: Overview of the design principles and their corresponding class [14].

2.7 Related Solutions

Related solutions are those that have already been developed, tested, and made available. These solutions may have been developed to identify a problem or need in the market and have created a solution to address it. The following examples can give valuable insights into techniques that concern our research questions. These solutions can be adapted or modified to meet the specific needs of a given situation, rather than starting from scratch.

2.7.1 Huskestue

Huskestue is a Norwegian game show that has been broadcast on TV-2 since 2015. The program is hosted by Kristian Ødegård and Dag Otto Lauritzen [29]. The show involves two competing teams, each of which is presented with questions from two primary categories over the course of each episode.

The show's design follows a blocked practice approach 2.2.2. Each category consists of three primary phases. During the initial phase, the contestants are asked easy questions that they are expected to know the answers to. In the second and main phase of each category, the contestants must participate in multiple games, with the winner of each game receiving the opportunity to answer a given question. If the selected contestant answers correctly, the opposing team typically receives some form of physical punishment, whereas incorrect answer results in the teammate of the answering contestant receiving the punishment. Finally, the contestants are presented with questions that they had previously answered incorrectly.

2.7.2 Kahoot!

Kahoot! is a digital quiz game that is extensively utilized in both educational and non-educational domains. It is an innovative platform that leverages game-based learning techniques to engage and entertain a significant number of players worldwide. Kahoot!'s broad user base extends across various settings, including schools, workplaces, and homes [11]. A typical game session involves multiple players joining a lobby, where the administrator starts the game. The game consists of several multiple-choice questions posed to the players. Points are scored based on answering quickly and correctly, while incorrect responses result in a deduction from the player's current score. To further add to the game's excitement, each question is timed, and failure to respond within the allocated time leads to a negative score. Upon completion of all questions, the game culminates with the presentation of a leaderboard that displays the top three players' scores. Kahoot! employs an interactive interface featuring music and colorful visuals that contribute to a playful atmosphere for the players.

From a literature review by Alf Inge Wang et.al they look at the effect of using Kahoot! and conclude that Kahoot! has a positive effect on learning performance, classroom dynamics, attuites, and anxiety [32]. However, the main challenge includes getting the difficulty level of questions and answers right. It is also mentioned that scoring based on how quickly the students answer reduces student reflection and causes some students to guess without thinking.

2.7.3 Quizlet

Quizlet is an online educational platform that provides a variety of study tools such as digital flashcards, quizzes, games, and practice tests, that are designed to aid memorization and improve understanding. It is mainly known for its use of flashcards which consist of a question on one side and an answer on the other side. In total, there are over 500 million user-generated flashcard sets and more than 60 million monthly users [23].

Tulce Kose et.al conducted a study on Turkish students learning English using Quizlet [13]. Initially, students studied and practiced target words without the online tool. After three weeks, they started covering the units they had previously learned with the use of Quizlet. The study concluded that Quizlet proved to be a valuable tool for repetition, particularly in the early stages of vocabulary, for students learning English as a foreign language.

2.7.4 Anki

Anki is an open-source software application that utilizes spaced repetition, as described in subsubsection 2.2.1, to enhance the retention of information [1]. In Anki, similar to Quizlet 2.7.3, users create digital flashcards that contain a set of questions and answers that they intend to memorize. The software then presents these flashcards to users at regular intervals based on their level of recall. Cards that are easily remembered are presented less frequently, while those that are challenging are presented more frequently making it particularly popular among individuals who need to memorize large quantities of information.

Anton Lambers et al. conducted a study on the impact of spaced repetition learning, utilizing Anki, among postgraduate trainees preparing for the Orthopedic basic sciences examination [15]. The study revealed that the students who dedicated more time to using a spaced repetition flashcard program, such as Anki, achieved higher scores in the examination. Consequently, the researchers concluded that employing Anki as a study strategy for achieving spaced repetition is highly effective, not limited to Orthopedic basic sciences, and can be applied to various other subject areas.

3 Design

The quality of design is important in order to effectively implement and test the techniques and ideas related to the research questions. It serves as a foundation for integrating and evaluating ideas and methods that address these questions, enabling us to convey essential information and generate meaningful insights.

RQ1: How can social interaction and computer-based learning be combined for an interactive QA system?
RQ2: How can question difficulty and order of questions be utilized?
RQ3: How can support for recall and memory retention be implemented?

We will design a quiz-based learning platform, where theory presented in section 2 will serve as a baseline for our incorporated mechanisms. This platform will provide a structured environment for interactive learning, allowing users to engage in vocal social interaction while benefiting from the advantages of computer-based learning.

3.1 Concept of Huskedu

In order to develop a suitable platform we created a mind map A which led to the concept of “Huskedu”. Huskedu is an innovative platform that aims to provide features and functionalities for user-generated learning materials, social interaction, and interactive quizzes.

One of the key features of Huskedu is the ability to engage in social interaction through open questions. Users will be able to interact, share knowledge, and learn from each other.

Incorporating elements of gamification and creating an enjoyable learning experience will enhance user engagement and motivation. Personalizing question difficulty will ensure that participants are appropriately challenged and motivated to progress.

To facilitate effective learning, Huskedu will implement support for recall and memory retention. Through repetition techniques, participants will reinforce their learning and improve long-term memory retention. Additionally, the platform will provide statistical feedback to participants, allowing them to track their progress and identify areas for improvement.

3.2 Choosing Game Design Principles

The theory from subsection 2.6 and Figure 4 will be used to tailor the proposed concept. Teemu H. Laine and Renny S. N. Lindberg indicate that these are meant to be helpful suggestions that designers can follow for educational game design problems at hand [14].

The design principle categories relevant to the proposed platform Huskedu are namely challenge, control, creativity, fairness, feedback, goals, learning, profile and ownership, relevance & relatedness, social play, and resources & economy. Figure 5 gives an overview of DPs that we found most important and relevant and can be included to increase motivation and engagement in the platform. The table includes a prioritization column ranging from high to low.

Category	**Design Principle**	**Importancy**
Challenge	DP 1: Provide challenges at various adjustable difficulty levels	**High**
Challenge	DP 2: Favor simple challenges over complex challenges	**High**
Challenge	DP 3: Provide enough time to solve challenges	**High**
Challenge	DP 5: Allow challenges to be repeated	**High**
Control	DP 10: Favor simple interactions	**High**
Creativity	DP 13: Allow players to create game content	**High**
Fairness	DP 20: Ensure similar opportunities of success regardless of experience	**High**
Feedback	DP 24: Provide access to performance data	**High**
Learning	DP 29: Provide cognitive challenges	**High**
Learning	DP 30: Embed assessment tools	**High**
Learning	DP 31: Choose time to reflect amid intense game moments	**High**
Profile & Ownership	DP 35: Offer past, present and future perspectives	**High**
Social play	DP 44: Provide opportunities for competition	**High**
Social play	DP 48: Allow players to distinguish themselves in their group.	**High**
Challenge	DP 4: Raise curiosity by interesting and/or unpredictable challenges	Medium
Control	DP 7: Use sensitive and accurate controls	Medium
Creativity	DP 14: Provide creative ways to solve challenges	Medium
Feedback	DP 21: Provide instructions and/or tutorials	Medium
Feedback	DP 22: Provide immediate, positive and useful feedback	Medium
Goals	DP 25: Create clear, meaningful, and achievable goals	Medium
Learning	DP 28: Provide relevant and pedagogically grounded learning content and activities	Medium
Profile & Ownership	DP 34: Provide the status of the game process and next available actions	Medium
Relevance and relatedness	DP 37: Relate to familiar activities	Medium
Relevance and relatedness	DP 38: Relate to past experiences	Medium
Resources & economy	DP 40: Enable collection of virtual goods	Medium
Resources & economy	DP 42: Introduce a possibility of losing resources	Medium
Social play	DP 45: Provide opportunities for collaboration	Medium
Learning	DP 31: Choose an appropriate transparency level for learning content	Low
Profile & Ownership	DP 33: Use a profile/avatar that the player can own and relate to	Low

Figure 5: Chosen design principles for Huskedu

DP 1: Provide challenges at various adjustable difficulty levels - According to the flow theory outlined in subsection 2.3, it is crucial to offer challenges at varying levels of difficulty that can be adjusted to accommodate diverse players' skill levels. This necessitates matching the difficulty of the questions to the player's knowledge may foster user motivation and engagement.

DP 2: Favor simple challenges over complex challenges - Overly complex and challenging questions can discourage players, leading to reduced engagement 2.3. To maintain user interest, it would be preferable to begin with simpler questions and gradually progress to more advanced ones in line with the player's development.

The same goes for the inclusion of minigames, where it is important to avoid overly complex challenges that potentially discourage users from playing.

DP 3: Provide enough time to solve challenges - In accordance with the self-explanation theory, subsubsection 2.2.2, it is advisable to give the player sufficient enough time to explain and rethink their definition. However, it should be a time limit to make the user feel constrained to give a concise and meaningful explanation without elaborating too much.

DP 5: Allow challenges to be repeated - By combining the theory of retrieval practice and the forgetting curve, as discussed in sections 2.1.2 and 2.2.2, the implementation of this design principle can serve as a important tool for reinforcing learning material and reviewing self-explanations.

DP 10: Favor simple interactions - In order to optimize the platform experience, it is crucial to ensure that the graphical user interface (GUI) is designed with user-friendliness in mind that aims to minimize the cognitive demands placed on players when attempting to understand the system. This design principle facilitates accurate task performance, enhances motivation, and enables effective engagement with the game's educational content.

DP 13: Allow players to create game content - In order for this platform to serve as an effective tool for enhancing a user's knowledge base, it is necessary that the learning content is dynamically expandable, equal to the user's level of competence. Furthermore, incorporating the possibility that allows players to engage in quizzes that others than themselves have created may serve to broaden the scope of their knowledge acquisition by exposing them to unexplored learning material.

DP 20: Ensure similar opportunities of success regardless of experience - Beginners and experienced users should be able to play games that is best suited for their competence. This will typically involve making degrees of difficulties for given categories, which will support motivation according to flow theory.

It is important that games are tailored to the competence levels of both novice and proficient users. Various degrees of difficulty within a quiz promote sustained motivation in accordance with flow theory 2.2.2.

DP 24: Provide access to performance data - Incorporating an interface that enables users to review their past performance on games has the potential to foster a desire for continued improvement. By leveraging insights from the forgetting curve, as outlined in subsubsection 2.1.2, the system can provide valuable data and increase a player's awareness of their retention rate.

DP 29: Provide cognitive challenges - The introduction of cognitive challenges can trigger cognitive curiosity, leading to the improved acquisition of knowledge and skills among players. This principle is strongly connected to both DP 1 and 13.

DP 30: Embed assessment tools - In an educational setting, an assessment system, as discussed in subsection 2.4, can serve as a valuable tool for educators to evaluate students' comprehension of a subject, such as a history course, and document their performance in a quantifiable manner. Additionally, it can be utilized as a self-monitoring mechanism for users to track their progress and obtain an overview of their academic achievements relative to a curriculum objective.

DP 31: Choose time to reflect amid intense game moments - Users may experience a high cognitive load when competing against others in the platform. Thus, it is important to provide breaks or "breathers" that can slow down the pace of the game.

DP 35: Offer past, present and future perspectives - Players should be provided with the ability to view their performance in both past and present games. An implementation of the forgetting curve could enable users to better understand the gradual decay of their memory of recently acquired knowledge. Effort and performance measurements may provoke the desire to achieve better.

DP 44: Provide opportunities for competition - Combining competition and collaboration can serve as a powerful means to enhance player motivation, by providing opportunities for gameplay against both themselves and others.

DP 48: Allow players to distinguish themselves in their group. - According to Teemu H. Laine and Renny S. N. Lindberg [14], allowing players to distinguish themselves within a group can demonstrate how individual contributions strengthen the team as a whole. Additionally, highlighting individual assessment alongside team assessment can promote the competence motivator by enabling players to learn about their personal strengths and weaknesses.

3.3 Playing Quizzes

As previously mentioned, one of the main aspects of the Huskedu concept is the inclusion of multiplayer quizzes. In order to ensure user engagement and motivation, it is essential to create a fun and playful game flow within the platform.

3.3.1 Utilizing Gamification

To optimize user engagement in the game flow we will utilize gamification to create a playful environment that enhances learning outcomes. Some strategies to achieve this will be to incorporate competition and repetition.

The combination of competition and learning can serve as a robust mechanism to motivate users. In the proposed solution, it would be relevant to incorporate points and leaderboards, within the context of social play and design principles 44 and 48. The integration of points and leaderboards has the potential to enhance user engagement and motivation, thereby increasing the effectiveness of the platform.

According to subsubsection 2.2.2, retrieval practice can improve the learning outcome. In the solution, it would be valuable to include previous questions as a means of testing a player's knowledge. This approach would encourage players to remain attentive throughout the game, as any question previously asked may reappear. Additionally, if a player answers a question incorrectly, they will be presented the correct solution, which can improve their memory recall. The repeated question requires a self-explanation to answer correctly, hence players may review their explanation and identify any gaps, errors, or new information, making it an effective platform of self-explanation, subsubsection 2.2.2. After completing the quiz, players will receive feedback on their performance and effort, enabling them to identify their strengths and weaknesses.

3.3.2 Game Flow

The quiz game will be divided into three main phases introduction, mini-games, and repetition. Each phase will focus on different aspects and are inspired and modified from Huskestue, subsubsection 2.7.1. In addition to the three main phases, we will follow up with statistics for users to gain more insight into their performance and effort.

Introduction - The purpose of the introduction phase is to determine the baseline knowledge of each player. During this phase, the questions asked will be relatively easy to answer. By doing this, an algorithm can adjust the difficulty level of subsequent questions to ensure that each player achieves a state of flow, as presented in subsection 2.3.

Mini-games - To create a playful and competitive environment between the players and teams, we include different mini-games. This is to increase in motivation, concentration, and engagement. When creating the mini-games it is important to vary in what type of skill is required to win a mini-game. This is to avoid rewarding the same people over and over for having a certain set of skills. In addition to this, answering the question incorrectly will lead to negative points awarded to the team in an attempt to ensure that the users priorities are also on the question and not just the mini-game.

Repetition - The purpose of the repetition phase is to induce retrieval practice subsubsection 2.2.2 among players, by recalling questions, answers, and solutions that have been given previously in the game. This may include questions that a player has answered incorrectly earlier in the game, as well as questions that other participants have answered.

Statistics - When the game is completed the players are taken to the leaderboard page, showing the winning team, followed by a personalized statistics page. This page is meant to illuminate possible knowledge gaps that can be improved further. A forgetting curve, subsubsection 2.1.2, will be displayed to make the user aware of memory loss and gain insight into the benefit of repetition.

3.4 Question Evaluation

The quiz game offers two question types: closed and open questions. A closed question can be answered with a short sentence or word, while an open question requires a more detailed response that encourages self-explanation. Although quizzes typically provide closed questions with or without multiple-choice, this solution aims to incorporate open questions that provoke discussions and align with self-explanation theory. There are three potential solutions for achieving this goal within the context of two teams, A and B, each comprised of two players, X and Y.

Multiple choice - In this scenario, each player will be presented with four options and prompted to select the correct answer. This format is commonly used in existing learning tools such as Kahoot subsubsection 2.7.2. However, implementing this format may limit the number of open-ended questions as the combination of multiple choice and open questions includes some challenges. To begin with, the player has to go through four distinct explanations, which can be time-consuming and challenging to compare. Additionally, the player cannot give their own self-explanation. Finally, the question maker would have to invest a significant amount of time in creating three false explanations for each question.

Game master - In this solution, a game master will be an independent role from the two teams, A and B. This person will be responsible for classifying answers as correct or incorrect, delegating questions, assigning points, and reading the questions. The platform will have a special interface designed for this role, which will allow the game master to take actions such as delegating points. The potential benefits of this game structure are that the game master can easily verify answers, include players, lead discussions, and give hints that can guide players in the right direction. However, this structure will require a minimum of three participants.

Voting system - In this game structure, each player has the option to vote for the correct or incorrect answer. After a player responds to a question, the corresponding solution is displayed and the other players can vote on the accuracy of the answer compared to the solution. This approach ensures that all players pay attention to the question and answer presented. If two or more participants disagree, they can start a discussion, potentially identifying errors or sharing new information. However, due to the competitive nature of the game, the voting system may be subject to bias.

3.5 Question Difficulty & Practice Sequencing

The sequencing of questions and categories presented to users can influence their learning outcomes. Two critical factors that can impact these outcomes are the level of difficulty of the questions and the degree of similarity among the categories. Understanding the effects of these parameters is essential for optimizing learning experiences and outcomes.

3.5.1 Question Difficulty

Based on the three flow theory states, subsection 2.3 and question difficulty subsubsection 2.5.1. (1) Fixed difficulty, can either lead to the boredom state or frustration state, depending on whether the difficulty is too high or too low. However, if the difficulty is just right, it will not lead to an increase in skill development or engagement, as the learner stays on the same difficulty. (2) The difficulty gradually increases, this leads to either of the three states. If the gradual increase in difficulty is too low, the learner can get bored, if its too steep, the learner can become frustrated, however if it by chance follows the users development, the user may experience the state of flow. (3) Adaptive difficulty, will be more likely to keep the user in a state of flow, as long as the adaption is done correctly. This requires the system to always try to challenge the learner, but not too much as it can lead to frustration and discouragement.

3.5.2 Adaptive Interleaved- & Blocked Practice

When developing the quiz platform it would be interesting to see to which degree interleaved- and blocked practice can be utilized to optimize the learning outcome. We assume, based on theory presented in subsubsection 2.2.2, that players learning outcome would be enhanced if questions are given in an interleaved order given two categories with a low to high degree of overlapping domain. Blocked practice would be beneficial in a context of two distinct categories. Interleaved practice will typically apply to categories such as physics and mathematics due to their positive correlation. It would also be desirable to implement this practice if two categories are two topics of the same superior category such as topics like division and multiplication, hence the superior category mathematics. On the other side, if the two chosen categories are totally independent of each other it would be desirable to use a blocked approach for questions due to the distinct domains. From the study by Svenja Heitmann et al. [9], the researchers tested if adaptive quizzing would provide higher learning benefits rather than non-adaptive. In the adaptive practice quiz questions were adapted to learners' expertise through cognitive demand ratings. In the non-adaptive condition quiz questions followed a fixed sequence. As a result, "[...] students in the adaptive practice quizzing condition outperformed those in the non-adaptive condition after a two-week delay, but not after a one-week delay. Exploratory mediation analyses show that performance on the quiz questions during the learning phase seems to be partly responsible for this effect." [9]. This will be interesting to look at in comparison to the forgetting curve, effort and performance.

3.6 Question Validity

With respect to the validity of questions, subsubsection 2.5.2 examines the validity of multiple-choice questions (MCQs). However, we aim to further investigate open-ended questions and assess whether any of the statistics derived for MCQs can be extended to open-ended questions, particularly the facility index. We presume that as MCQs provide the opportunity to arrive at the correct answer by guessing among the provided alternatives, the discrimination efficiency index is a more realistic measure. In contrast, in open-ended questions where the answer options are not provided, guessing the correct response is more challenging, rendering the discrimination efficiency index less relevant. Nevertheless, the facility index can be useful in determining whether a question is too difficult or poorly formulated.

3.7 Visual Design

When creating the platform we want to separate between a playful environment and a more serious environment based on where the users are on the platform.

With the use of colors, fonts, emojis, and avatars we want users that are in the process of playing a game to experience a playful environment. Here the shape s and fonts are more rounded, and the colors are brighter and have high contrasts. We also want to include toon-based avatars to emphasize playfulness.

In more administrative steps such as creating a quiz, we want to remove the playful environment to ensure that the users take this process seriously. This is to avoid having quizzes with incorrect solutions and quizzes with unserious questions. Here we will be experimenting with less contrasts, fonts that are more commonly used in school, papers and articles.

4 Implementation

In this chapter, we will introduce the implementation of our proposed quiz game platform. This includes the development of a vocal answering format for both closed and open questions, the design of an algorithm for optimized question order, and the creation of a playful environment with balanced mini-games to achieve the flow state. Additionally, we will explore the potential of our platform as a supplementary learning tool, providing users with a new variation of how they can study with others.

4.1 Prototyping

To identify potential problems, limitations, and areas for improvement, it is crucial to conduct sufficient testing. User testing is an effective means of gathering valuable insights into how potential users perceive an application. By comparing multiple prototypes at varying fidelity levels, one can determine the most effective and feasible design and make the technical implementation process more efficient. Therefore, in this study, we developed multiple prototypes to aid in identifying potential issues and improvements in our platform.

The prototypes were tested by introducing a use case scenario for the users, followed by some pretesting questions to get an understanding of their relationship to learning, gamification, repetition, and preferred learning type. We also had some post-test questions to get a deeper understanding of their thoughts on the concept.

4.1.1 Low Fidelity Prototype

Our low-fidelity prototype consisted of a paper prototype printed and sketched on an iPad, Figure 6. Before settling on a version that was suitable for testing, we iterated over this prototype several times. From the user tests' we discovered that certain concepts were unfamiliar to users and required additional explanations. Additionally, we found that certain parts of the game flow were difficult to follow, requiring introductions to provide insight into what would occur next.

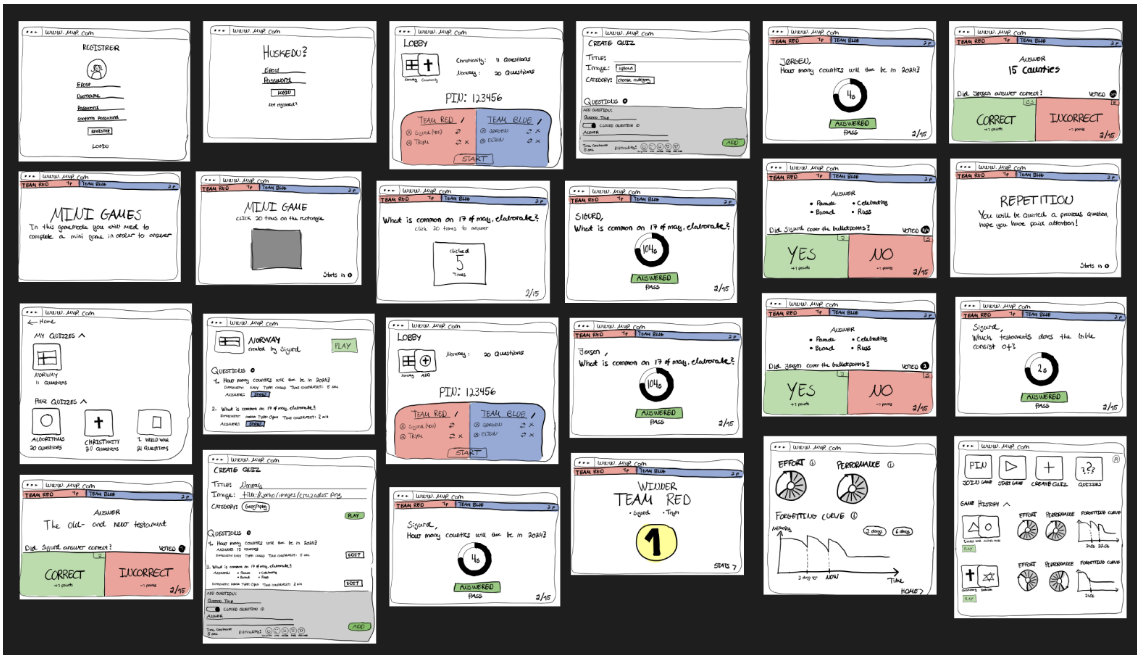


Figure 6: Overview of the final screens produced on the iPad

4.1.2 High Fidelity Prototype

The high-fidelity prototype, which was developed using Figma, represented a more advanced version of the paper prototype, Figure 7. It was a highly refined prototype that allowed users to provide feedback on both the concept and the design itself. Users provided feedback on the design that highlighted some minor issues but were important for improving the user experience. These included maintaining a playful environment throughout specific pages to avoid confusion as well as addressing minor issues related to buttons and user navigation in our final product.

Furthermore, testing with the high-fidelity prototype provided valuable insights into the game flow, such as the need for more specific introductions and clarifying whose turn it was to answer. Additionally, the testing revealed that some users found the concept of using a digital tool combined with social interaction to be unfamiliar. Therefore, we had to be precise in the language used throughout the application to avoid any confusion.

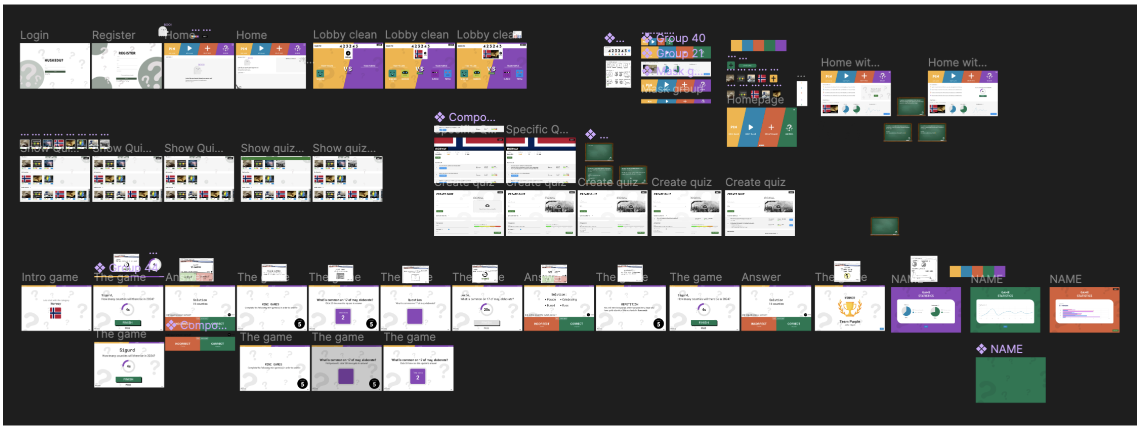


Figure 7: Overview of the final screens produced with Figma

4.2 Technical Implementation

Huskedu is built with React, Typescript, and Firebase. The use of these technologies offers advantages, such as high performance, enhanced functionality, and efficient data storage and management. The Firebase platform provides several services that are utilized in the development of Huskedu. These services include Firebase Firestore, which is a flexible and scalable document database that enables applying changes and extending functionality with ease, thus facilitating an agile development process. Firebase Storage is utilized to store images that are linked to the database. Furthermore, Firebase Hosting provides a straightforward method for deploying the website, making it easily accessible to actual users.

Moreover, to ensure seamless and secure authentication for users, the registration process is made available through Firebase Authentication Single sign-on (SSO), which is restricted to Google. By utilizing Firebase, there is no requirement for server setup, enabling the initiation of the application development process without delay. Finally, Firebase Cloud Functions are employed to listen for updates in the database and to act accordingly in response to changes. This integrated use of Firebase services enables Huskedu to function efficiently and effectively, providing a comprehensive educational platform for its users.

4.3 Features Implemented

This section offers a detailed overview of the Huskedu platform, including its main use cases illustrated in Figure 8. The development of the platform was the result of extensive user testing, which incorporated feedback from both the low-fidelity paper prototype 4.1.1, and the high-fidelity prototype 4.1.2. Through an iterative design process, the platform's functionality was refined based on user feedback, resulting in an enhanced user experience.

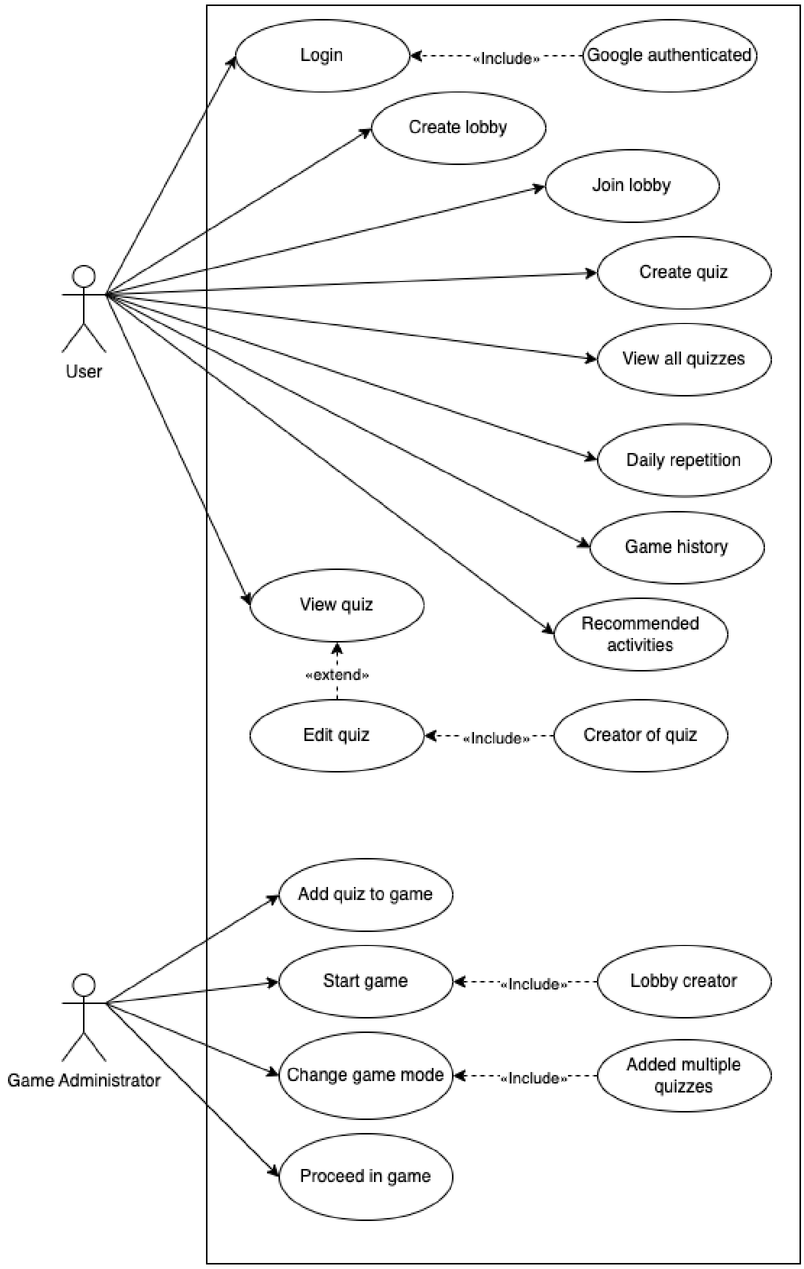


Figure 8: Use case diagram illustrating the main functionalities of the platform

4.3.1 Login

Upon initial access to the platform, users are prompted to either login or create an account using their Google credentials, Figure 9.

Figure 9: Login screen

4.3.2 Home Screen

After logging in, the main screen Figure 11 is presented to users, providing access to all available main functionalities. The use of vibrant colors is intended to create an engaging and playful environment. For users who have not yet participated in any gaming activities, a different version of the home screen is displayed, Figure 10. However, for those who have previously played games on the platform, an overview of their match history, recommendations, and daily challenges will be displayed, Figure 11. The top navigation bar presents four options: "Join Game", "Host Game", "Create Quiz", and "Quizzes", representing the core functionalities of the platform.

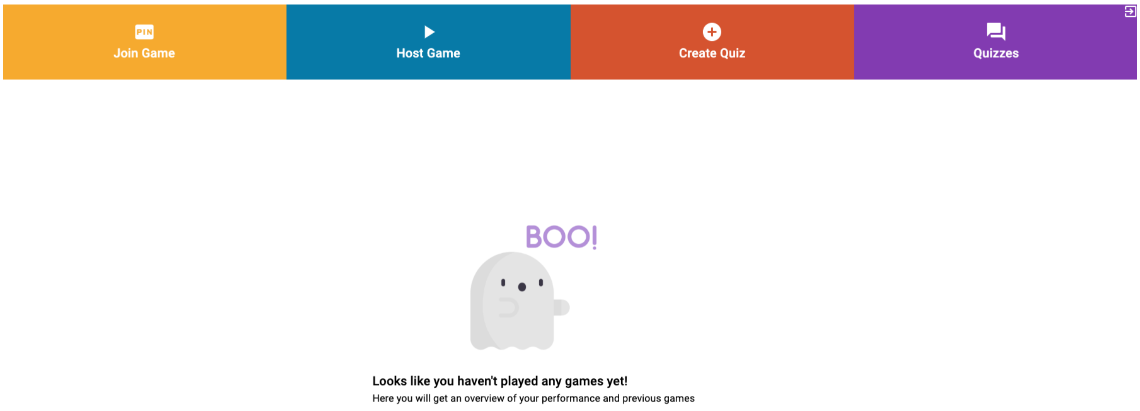


Figure 10: Home screen for user without any played games

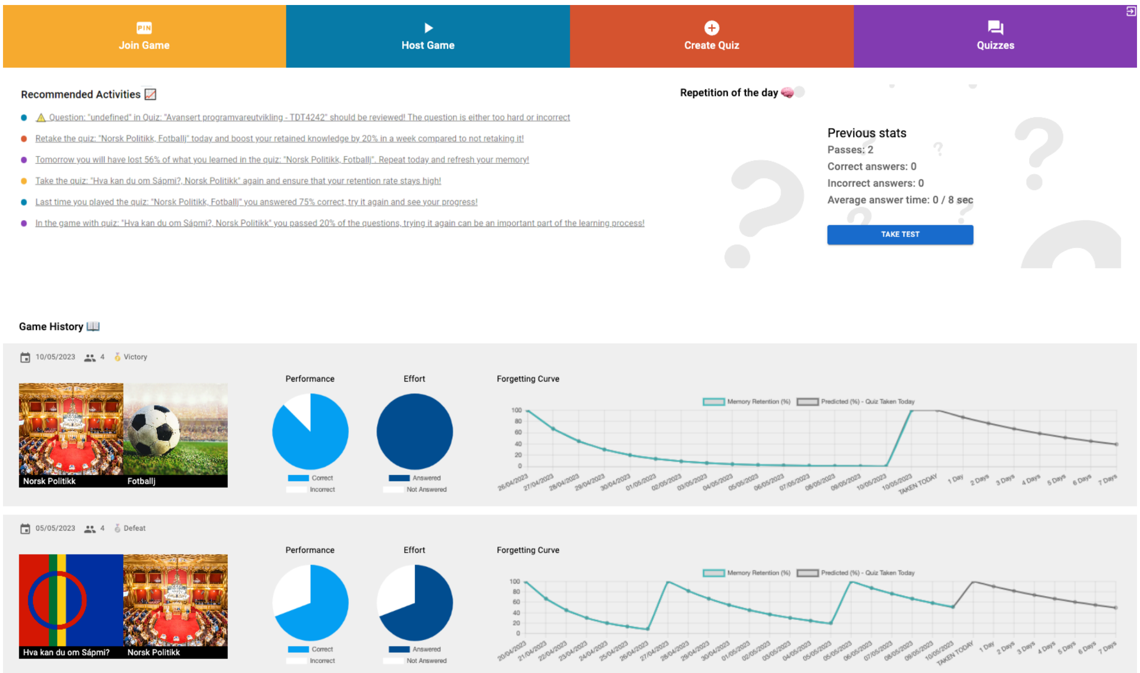


Figure 11: Home screen with game history, recommendations and daily challenge

4.3.3 Create Quiz

In order to create a quiz, users are required to provide a title, select a category, and upload a cover image. Upon completion of these initial steps, users can proceed to add questions to the quiz. When adding a question, it is necessary to include the question itself, along with the answer and additional information, such as an estimate of difficulty and a time constraint for answering. When giving an estimate of the difficulty users are provided with a five-point Likert scale, ranging from very easy to very hard. A Likert scale is a rating scale where respondents choose to indicate their opinions, attitudes, or feelings about a particular issue [19].

The graphical user interface for creating a quiz is illustrated in Figure 12.

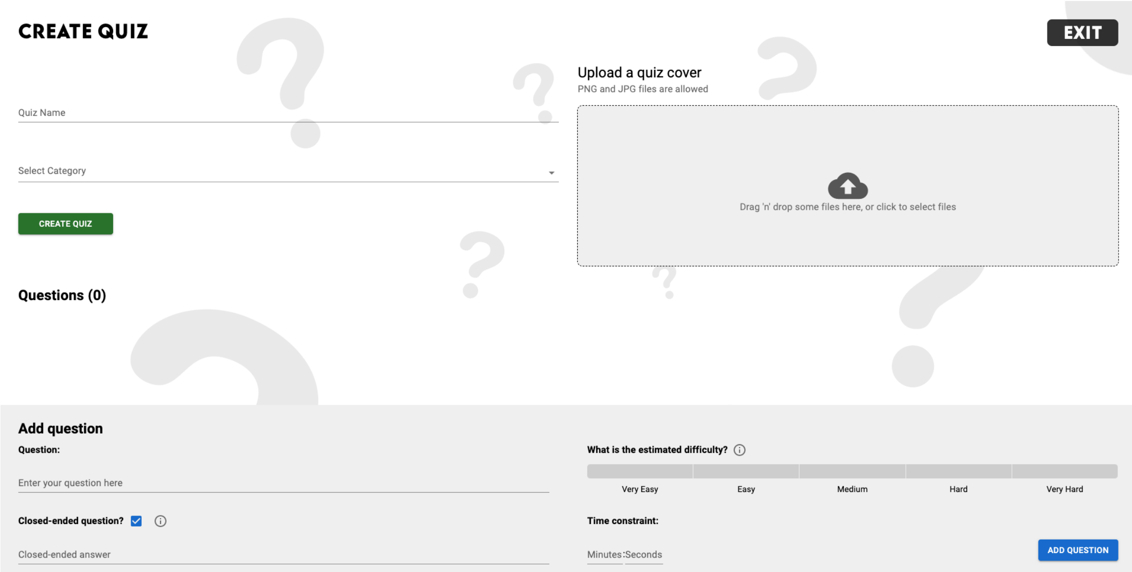
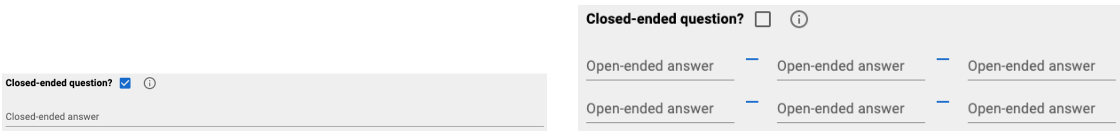


Figure 12: The page for creating a quiz and adding corresponding questions

Furthermore, users have the option to add either closed- or open questions. This can be chosen by using the checkbox above the designated field of inputting the answer. By choosing closed questions, the quiz creator has one input field for their answer, Figure 13(a). Choosing an open question, the creator has the possibility of adding and removing bullet points for the players to cover during the game, Figure 13(b).



(a) Closed question

(b) Open question

Figure 13: How closed and open question answers are added through the creation of a quiz

4.3.4 Quizzes & Viewable Quiz

Navigating to “Quizzes”, users are presented with an overview of both self-created and other user-generated quizzes, as depicted in Figure 14. Each quiz is represented through an image cover, title, number of questions, and an average difficulty rating. Should users require further details on a specific quiz, they can navigate to the dedicated information page by clicking the ”View” button.

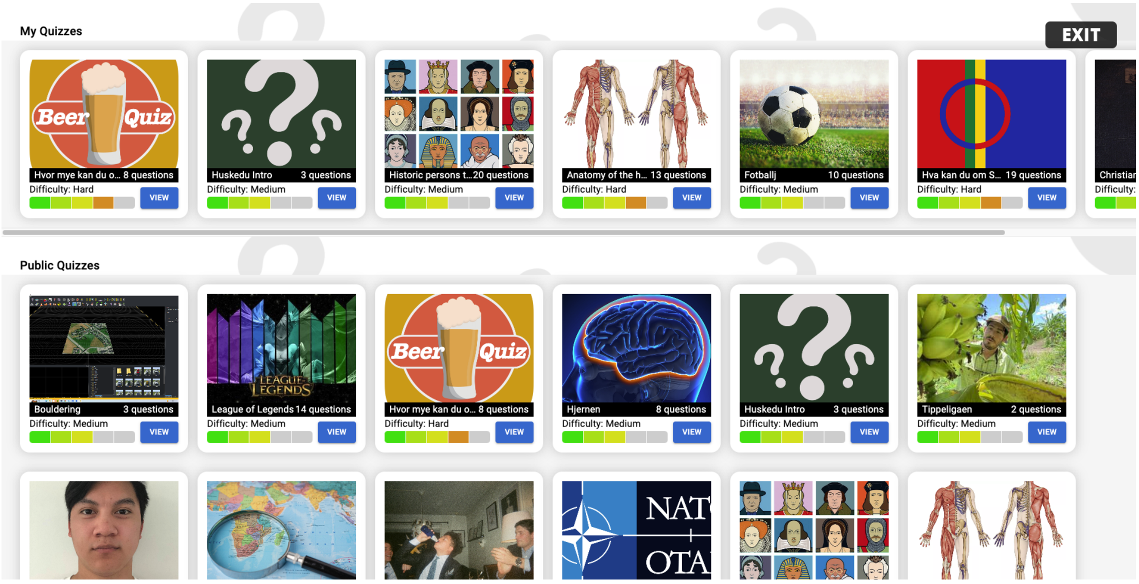


Figure 14: Overview of user-generated quizzes

When viewing a specific quiz, the page presents a comprehensive overview of the quiz content. This includes all the questions and their corresponding answers, which can be accessed by selecting the “Show Answer” option, Figure 15. This design feature is intended to spark interest and motivation by enabling users to review and assess the questions before incorporating the quiz into a game, without seeing the solutions and gaining an unfair advantage when playing the quiz.

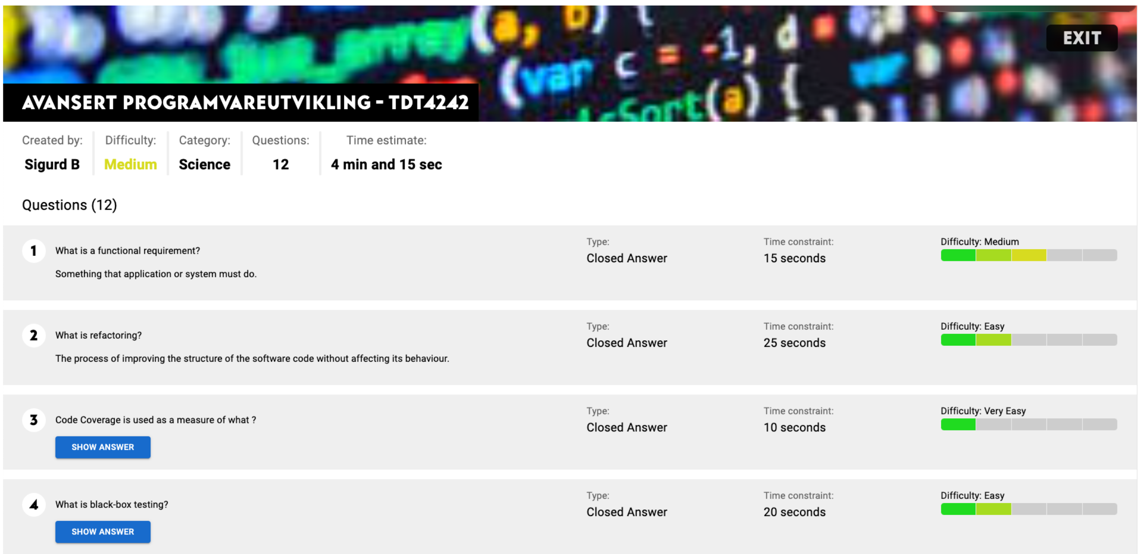


Figure 15: Information page for the quiz ”Avansert programvareutvikling - TDT4242”

4.3.5 Helper Modals

As stated in subsection 4.1, the platform includes several terms that are not commonly understood by users, while being important for the overall experience. In order to address this issue, we implemented a solution that involves adding a universal info button next to unfamiliar terms that impact the user experience. Clicking on this button opens a modal with help text displayed on a blackboard image, aimed at clarifying the concerned term. Figure 16 provides an example of this functionality, where the user is presented with the option to choose between blocked and interleaved practice when two or more quizzes are available in the lobby. Clicking on this button prompts a blackboard modal displaying an explanation of the terms. The modal can be closed by clicking anywhere outside of the blackboard.

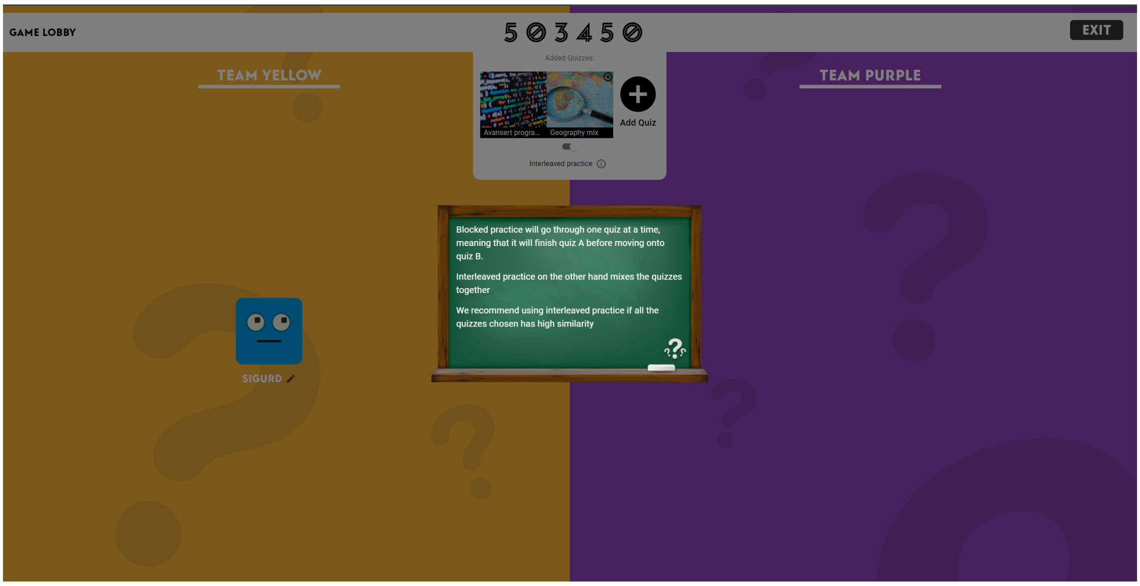


Figure 16: One of the helping popups explaining uncommon terminology

4.3.6 Lobby & Game

In order to participate in a game, users are required to join a lobby, a process which can be initiated either by accessing the "Join Game" option, as elaborated upon in subsubsection 4.3.7, or by selecting "Create Game". As depicted in Figure 18, the graphical user interface (GUI) will adjust accordingly. Once users have successfully joined a lobby, they will be automatically allocated to different teams. To enhance the playful and engaging atmosphere, each user is assigned a randomly generated avatar. The game setup and logic are further described in subsection 4.4.

4.3.7 Join Lobby

Each lobby has a unique pin code, serving as the exclusive method of accessing a lobby. In most cases, Huskedu is intended for use in physical gatherings, in which the pin code can be shared verbally or through a messaging service. Once in possession of the pin code, users must input it into the pop-up, Figure 17, that appears upon clicking on the "Join Lobby" tab.

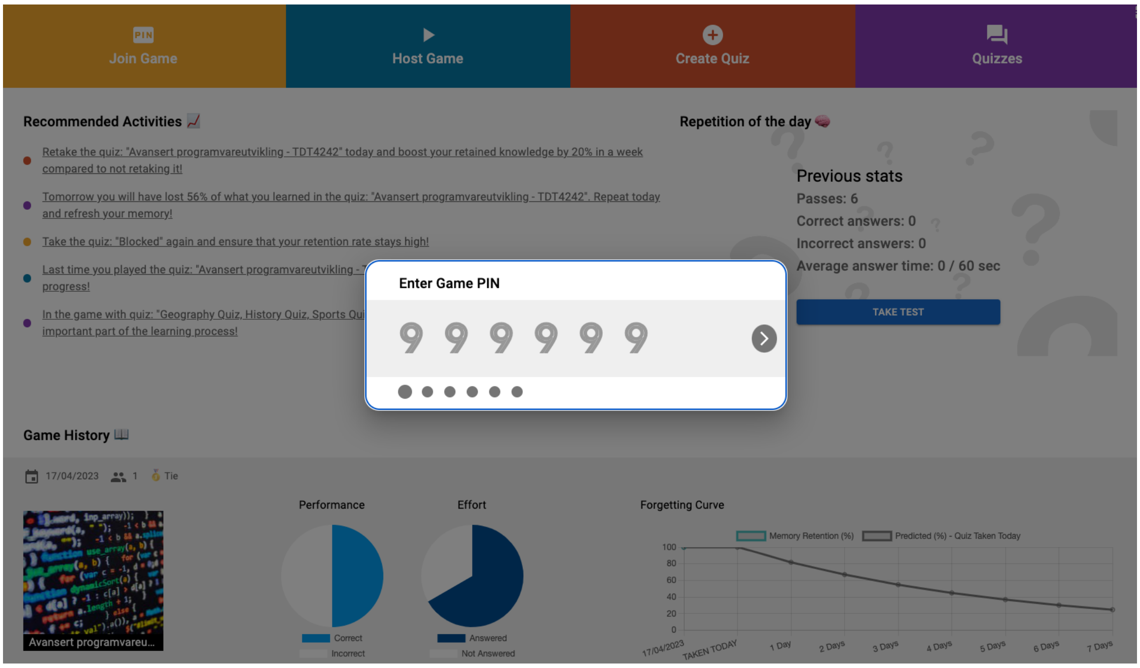


Figure 17: The modal pop-up for inputting the given pin code

4.3.8 Lobby

To initiate a game, a user needs to create a lobby that others can join. Once the lobby is created, Figure 18, the user becomes the administrator who is responsible for adding quizzes, adjusting game settings such as interleaved- or blocked practice, and controls the pace of the game. Once all players are present and ready, the administrator can start the game.

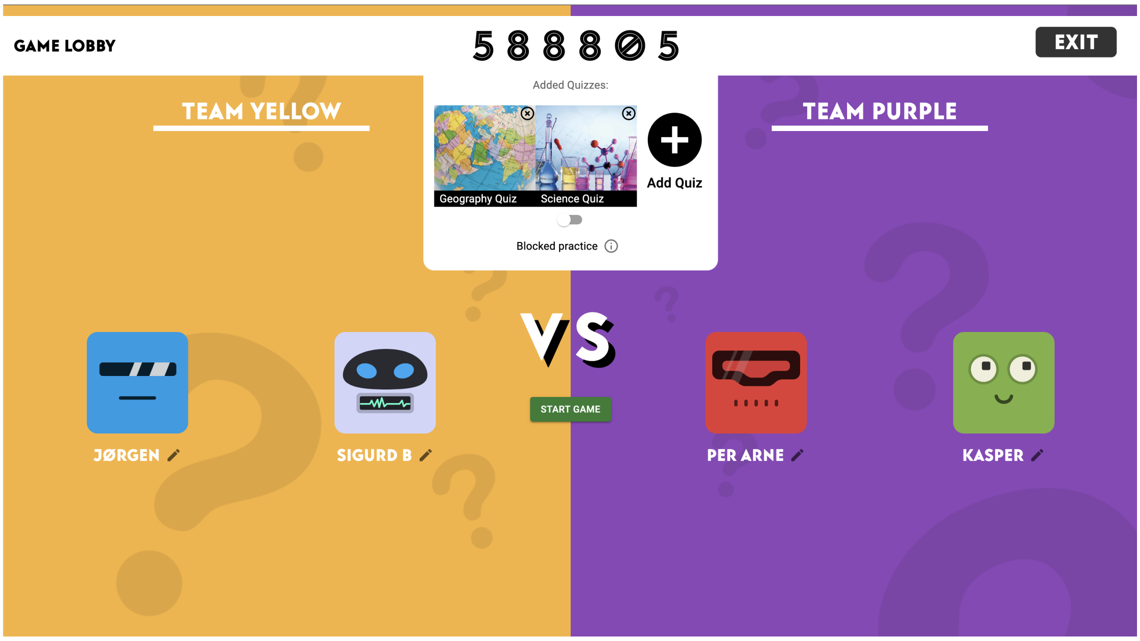


Figure 18: The lobby for others to join. This is from an administrator's perspective.

4.4 Game Flow

The game is divided into three phases, the introduction, mini-games, and repetition. The introduction phase is responsible for engaging all players in the quiz by asking questions, while in contrast, the first player to complete the mini-games will get the opportunity to answer the question. These mini-games serve as both a break and an engaging part of the game that aims to increase player engagement. Towards the end of the game, players will be asked repetition questions, which are questions that have been answered earlier, in order to support recall from memory. The game concludes with a scoreboard displaying the winner. Players will have the option to view additional insights about their performance, effort, and future memory retention.

The logic of the game is illustrated as a diagram in Figure 19. For each new phase or category, the platform will show an introductory page to inform the user. The question may be assigned to a specific individual (introduction- and repetition phase) or it can be a free-for-all question where players compete in a mini-game to earn the right to answer. After a question has been answered, users vote to determine the correctness and proceed to the next question. This flow is repeated until all questions in the three phases have been queried and answered.

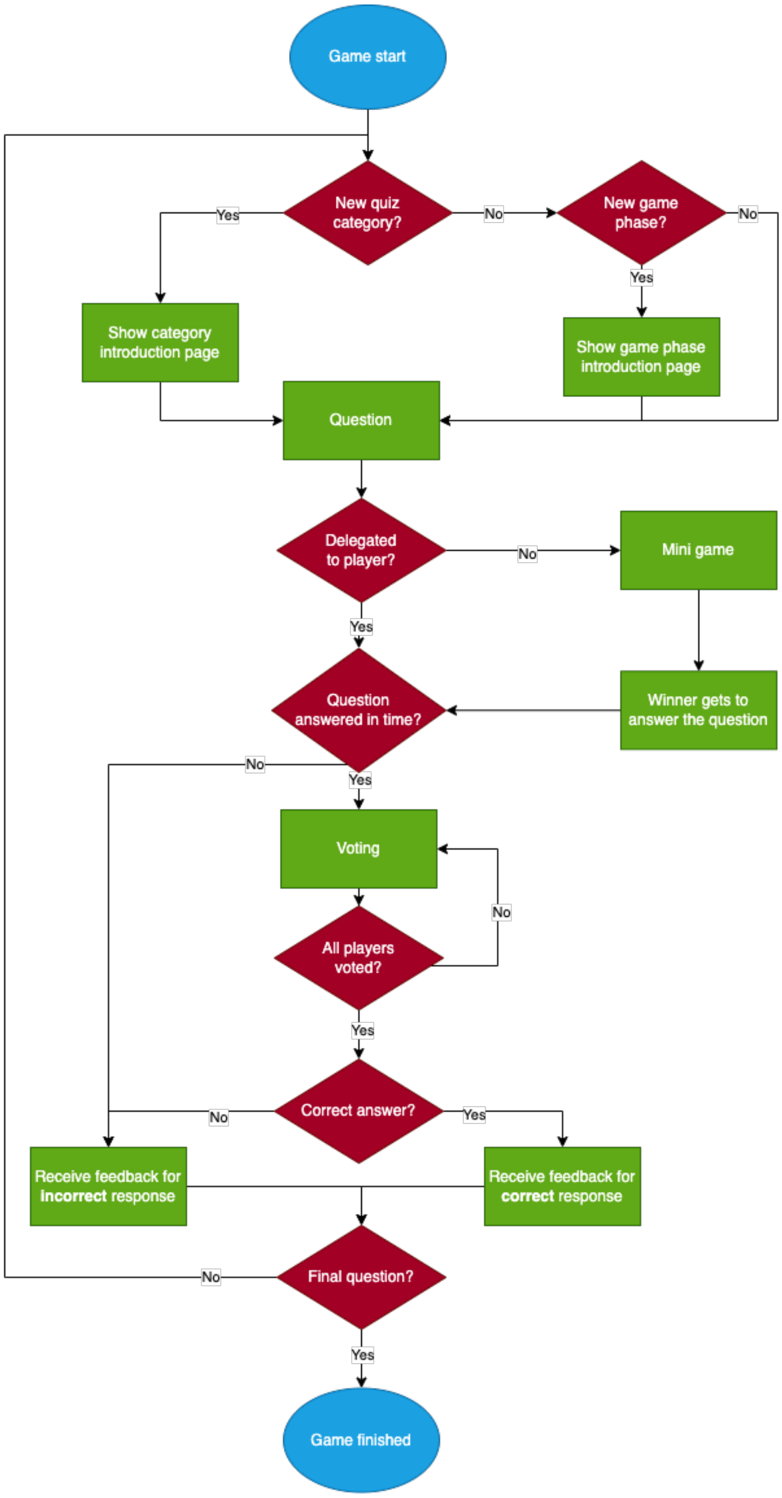


Figure 19: Diagram showing the flow of the game

4.4.1 Distribution of Intro-, Mini-games-, & Repetition Questions

Since the game consists of potentially multiple players and quizzes, it is important to distribute the number of questions equally to ensure fairness. Therefore, We have developed an algorithm that handles the distribution of questions to the three different game phases. As the number of questions and the number of players can vary, the algorithm finds the closest distribution to 2/3 of the quiz's questions while ensuring that each player gets the same amount of questions from each quiz included, for the introduction phase. The remaining questions which are approximately 1/3 can both be even and uneven as it is distributed as minigames, where all the players have the opportunity to answer the question.

When adding quizzes, the algorithm includes all questions from all quizzes and uses the three formulas to distribute the questions for each phase. The three formulas are I(x,y), G(x,y) and R(x,y), where x represents the number of participants and y is the number of total questions.

I(x,y) is used to find the number of questions included in the introduction phase:

$$I(x,y) = \lceil \frac{\lfloor 2y/3 \rfloor}{x} \rceil \cdot x$$

G(x,y) takes use of I(x,y) and is used to find the remaining questions for the minigame phase:

$$G(x,y) = y - I(x,y)$$

To determine the number of repetition questions, we aim for approximately half of the quiz questions, with each player receiving the same number of questions to ensure fairness. The number of questions included in the repetition phase is found by using R(x,y). If there are multiple quizzes, this formula is used individually for each quiz, to ensure that it takes approximately half of each quiz and not half of the total amount of questions in the game.

$$R(x,y) = \lceil \frac{\lfloor y/x \rfloor}{2} \rceil \cdot x$$

This algorithm ensures a fair game for all players, regardless of the number of players or quiz questions involved.

4.4.2 Blocked & Interleaved Mode

Within the lobby, the administrator holds the ability to modify the dynamic flow of the game by activating either the blocked or interleaved practice. In the event that the blocked practice is selected, each quiz and its corresponding questions will be presented separately, as outlined in subsubsection 2.2.2. This entails that each quiz will commence with introduction questions, followed by mini-games, and then proceed to the subsequent quiz.

On the other hand, if the interleaved practice is enabled, the game flow is split into these three phases, but with questions from different categories mixed together. The two practices both include a repetition phase after all questions are queried. These two practices are further demonstrated in Figure 20.

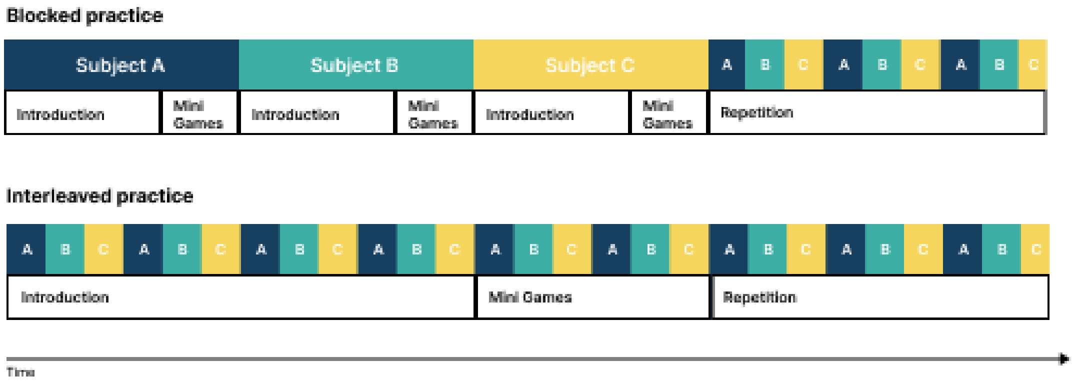
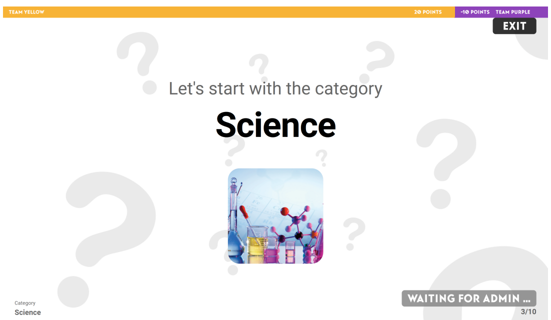


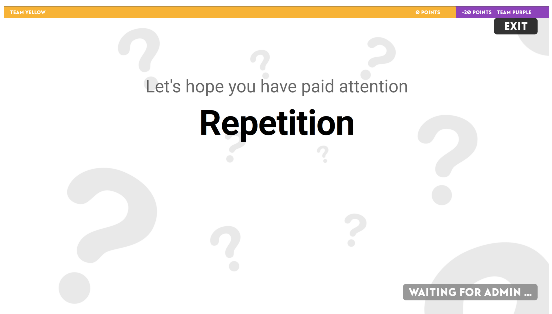
Figure 20: Illustration of interleaved and blocked practice is combined with the game flow of intro-, mini-games and repetition questions.

4.4.3 Information Pages

Whenever the game state changes between phases or categories the users are met with an information page. This is to give the users a brief pause of effect to let them absorb the information given to them before moving on to a new phase or category of the game, and to minimize confusion.



(a) Example category change

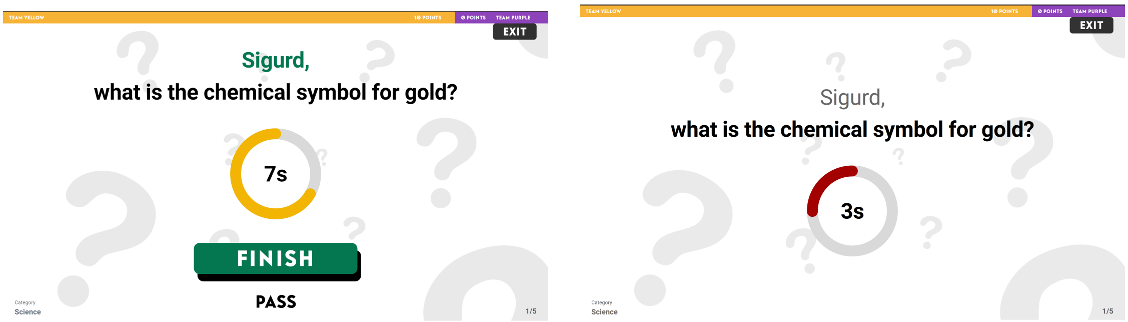


(b) Example of phase change

Figure 21: Examples of category and phase change

4.4.4 Querying

Whenever a user is queried a question the system will highlight their name and display buttons for finishing the question or passing the question as long as the question is delegated to them. Otherwise, the name will have a default color and no buttons will be displayed Figure 28. The timer of the question will be the same for all players in the game. When queried a question, the user must answer vocally within the timer and press the finish button to proceed to voting.



(a) Example of being queried a question

(b) Example of other players being queried a question

Figure 22: The two different views of a question being queried

If the user presses pass or does not press finish within the time constraint, the system will reveal the solution and leave room for the users to take in the information presented.

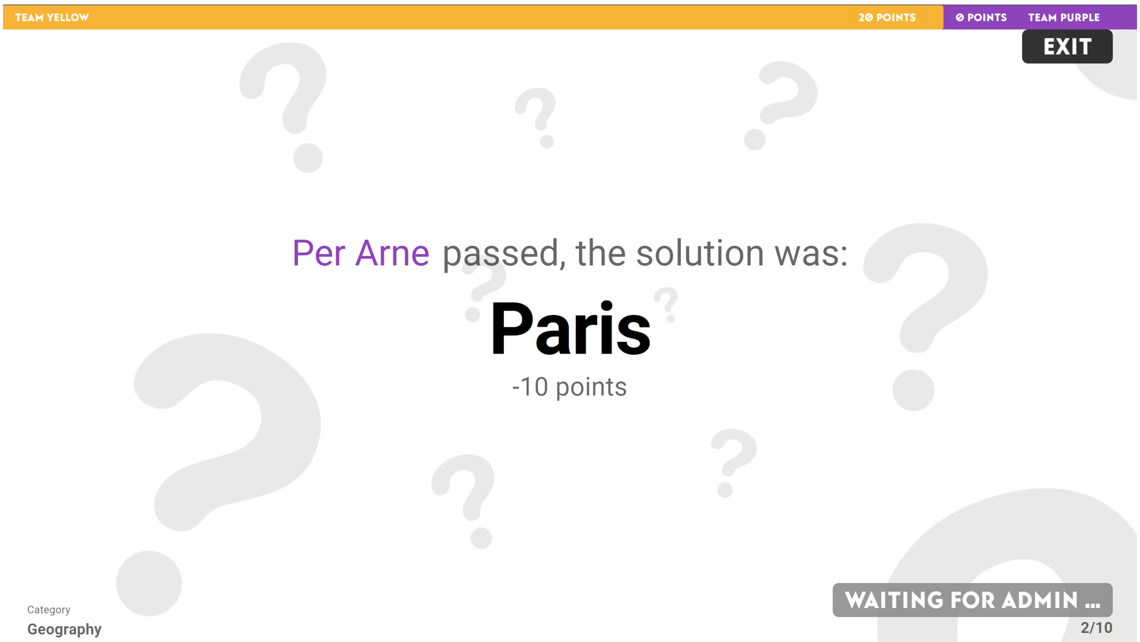


Figure 23: Information screen when a question is passed

4.4.5 Answer Validation

As discussed in subsection 3.4, we considered multiple options on how to validate answers given by the users. As our goal is to have more open questions that can lead to discussions and self-explanation we opted for the solution of a voting system. This way we have our delegation algorithm assigns a fitting question to the user, who then has to answer vocally before the timer runs out. Once the user presses finish, the system reveals the solution, and the other users must discuss and vote on whether the answer by the user compared to the solution is correct. The answer is defined as correct or incorrect based on the majority of player votes. In the case of a tie, the system considers the user's answer to be correct. In Figure 24 we show an example of the solution being presented and the voting system in action, while in Figure 27 we show the votes and what it results in.

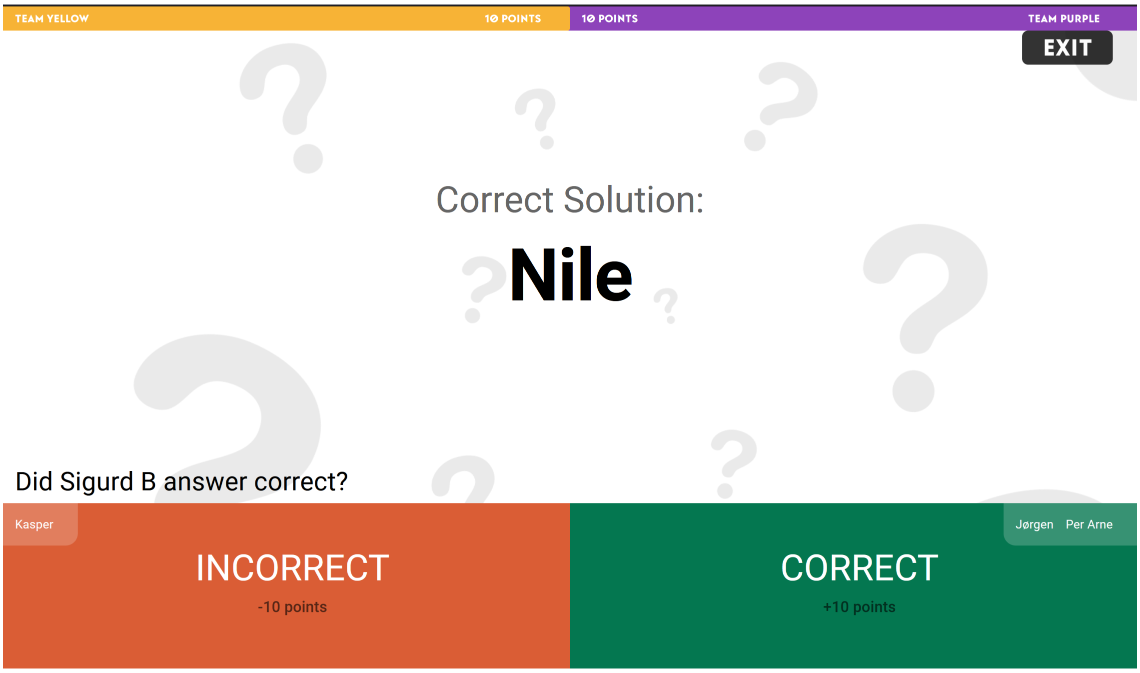


Figure 24: Voting system where the current votes are displayed

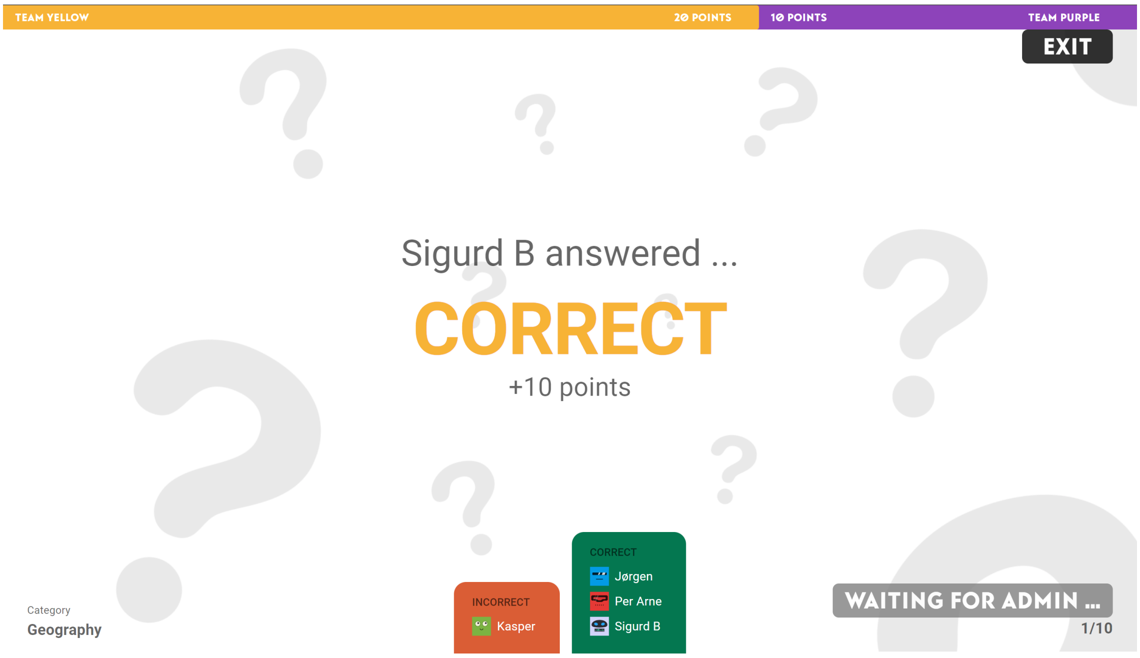
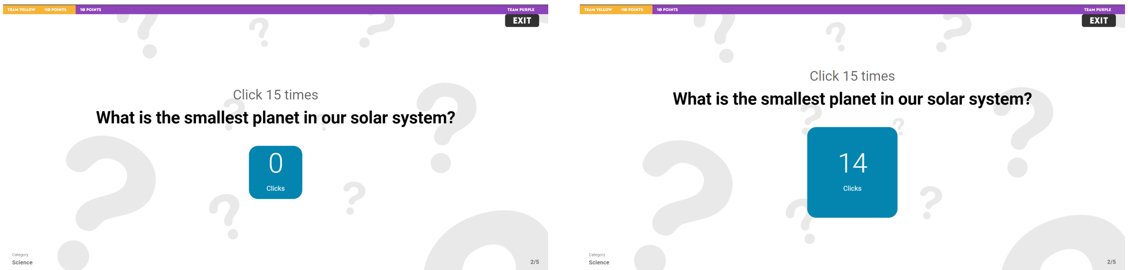


Figure 25: After the vote goes through

4.4.6 Mini-games

As discussed in subsubsection 3.3.2 we talk about how it is important to vary in what skillset is favored for winning a mini-game, therefore, we came up with three different mini-games with various skillsets in mind. The first one is inspired by attempting to pop a balloon by filling it with too much air. Our digital solution to recreating this was to click a square 15 amount of times to make it pop, where the first person to make it pop gets to answer. The square itself grows larger with each click. This mini-game in particular focuses on rewarding speed. We moved from having a balloon sprite as the objective to a square as this had better visibility and accessibility.



(a) Square at zero clicks (b) Square at 14 clicks

Figure 26: Mini game: Clicking the box

The second mini-game favors coordination and reaction. This game is inspired by whack-a-mole, a traditional game often available at amusement parks. The digital solution to this was that we created a set amount of boxes where the mole would pop out of, where the users have to click the mole. The first one to whack the mole 15 times gets to answer the question. For each click on the mole their current amount of clicks increments by one.

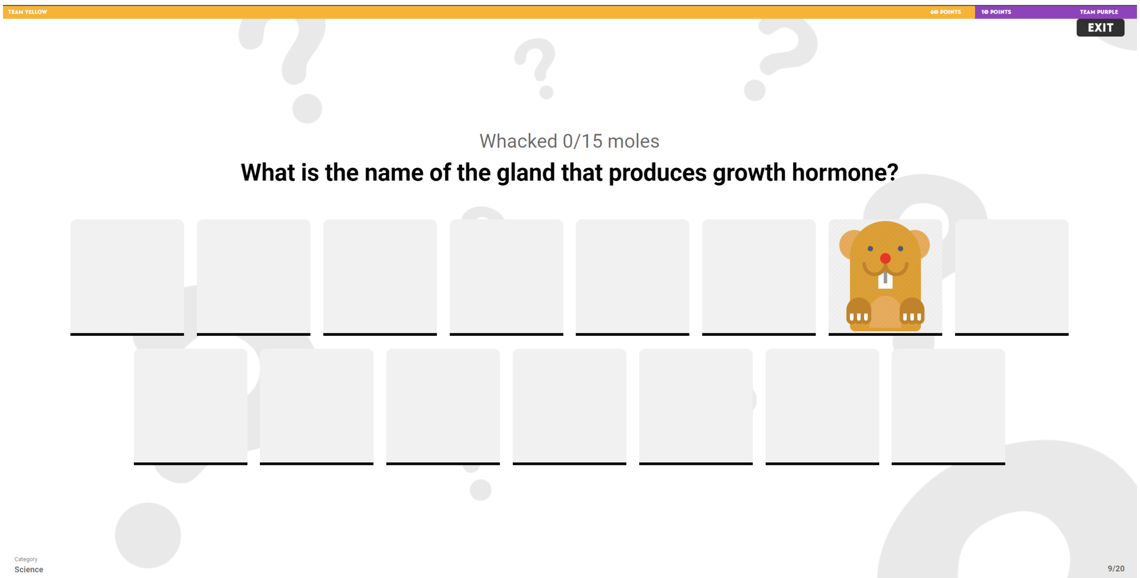
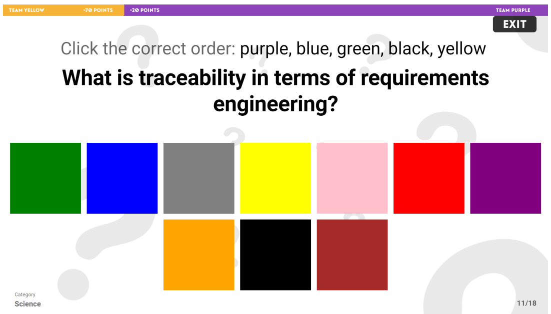
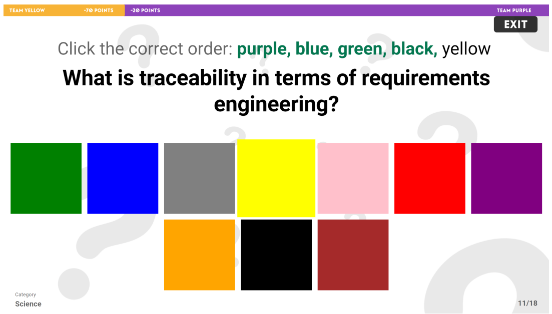


Figure 27: Minigame: Whack-a-mole

The third mini-game favors focus and sequencing, as the mini-game involves the ability to mentally organize information and follow a logical progression of steps while being fast. The game displays 10 squares with different colors, and the users get a list of 5 colors that must be clicked in a specific order to be able to answer the question. As the user clicks one of the colors the instruction text will change text color from default to green to give the user feedback on their progress.



(a) Zero colored boxes clicked



(b) Four colored boxes clicked

Figure 28: Mini game: Clicking the colored box in given order

Several other mini-games were discussed and prototyped but eventually withdrawn as they did not fulfill the requirements of being user-friendly, fair, and introduce a new skillset to favor.

4.4.7 Repetition

In the final stage of the game, there is a repetition phase. Here the users get to answer questions they have answered incorrectly or questions queried to other players. The questions assigned to a player are determined by their performance in the current game. An algorithm scrutinizes a player's game history and prioritizes the following:

First, all answered questions are separated. Then, the algorithm arranges all answers that were not given during the repetition phase. Next, it checks for incorrect answers provided by the current player. If this list is empty, it searches for incorrect answers given by other players. If that list is also empty, it examines correct answers provided by other players. Finally, if none of the above criteria is met, the algorithm provides a question that the player has previously answered correctly, which only happens if a player is playing alone.

4.4.8 Allocation of Game Points

To increase the competition between teams there was implemented a scoring system. A team is allocated 10 points if the person on the corresponding team answers correctly. If the team player answers incorrectly or passes they will be subtracted

10 points. In the case of an even distribution of incorrect/correct votes, the team will receive 10 points.

4.4.9 Implementation of Game Logic

One of the technical difficulties encountered during the development of the game was the need to ensure real-time updates for all participants. To achieve this, Firebase real-time listeners were utilized, simulating a socket connection in which each connected device would listen for any changes in the game's database. The data stored in the database is structured to indicate the current phase of the game and which participant is responsible for answering the current question.

The website's logic handles this data by mapping it from the database to a game controller on the frontend, thereby displaying the appropriate content to the user. Instead of relying on the backend to update the logic and delegate the correct question, the logic and delegation is calculated on the administrator's device. The reason for this solution is to let the administrator control the pace of the game.

the platform verifies whether the player is an administrator, granting them the ability to proceed to the next question.

When the administrator proceeds, a function is triggered on their device to handle the game's logic. This function compares the current game state to either an interleaved or blocked practice, identifying and delegating the optimal question based on the player's game history. The selected question is then posted to the database and shared with the other players.

It is worth noting that the administrator's content will not be updated ahead of others, as the database listener only checks for updates in the database and subsequently updates the graphical user interface. However, a potential disadvantage or advantage that may arise is if users are connected to different networks with varying bandwidths. This discrepancy may result in some users receiving GUI updates before others, particularly advantageous during the mini-game phase where completion of the game is necessary to provide an answer and earn points.

4.5 Question Difficulty

Rather than choosing quiz questions randomly, an algorithm is employed to determine the ideal difficulty level based on a player's prior performance to sustain a state of flow. Using a five-point Likert scale, **??**, users are asked to give an estimate of the difficulty of each question, which can then be utilized as an parameter in the algorithm. The adjustment of question difficulty is determined through quantifiable data and automated to ensure consistency.

4.5.1 Estimating & Adjusting Difficulties

In the process of creating a quiz with multiple questions, users are required to provide an estimate of the difficulty level for each question. However, this approach is largely subjective and influenced by the user's own perception and knowledge. As a solution, a scheduled cloud function is implemented to calculate the facility index for all questions in the created quizzes on a weekly basis, based on the facility index from subsubsection 2.5.2. This function identifies the first instance when a distinct player answers a question and calculates the facility index accordingly. If the calculated facility index differs from the current difficulty estimate, the function updates the question difficulty level. This approach ensures that the questions' difficulty levels are adjusted based on objective measures rather than subjective opinions. If the facility index is below the critical level of 5% the quiz creator will be prompted to revise the specific question.

4.5.2 Finding the Ideal Difficulty

As discussed in section 2.3, the flow state is a desirable outcome in the quiz, whereby players are optimally challenged by questions that match their level of knowledge and performance. To achieve this, an algorithm has been developed that tailors the difficulty of questions based on the player's previous performance. Specifically, the algorithm compares a player's past answers and assigns questions that are appropriately challenging. If a player has answered previous questions correctly, they will be given a question of higher difficulty. If the player answers the question incorrectly they will receive a question of the same difficulty. However, if the player answers incorrectly two times in a row, the player will be given an easier question. When a game initially starts all players are assigned the same difficulty level, which is set to Medium. The algorithm for finding the difficulty is further illustrated in Figure 29.

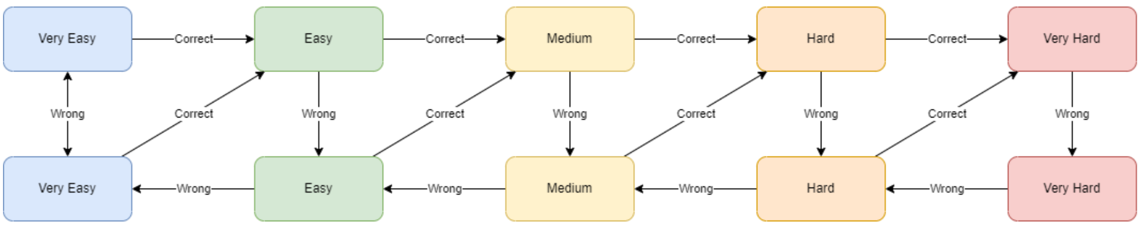


Figure 29: Illustration of how ideal question difficulties are delegated

4.6 Daily Challenge

In addition to the repetition within the game flow, we have implemented an algorithm that selects the most challenging question for the user to review based on their game history. This selection is not based on questions' difficulty level, but rather on the user's previous performance on questions. Specifically, we identify the question that the user has answered incorrectly or skipped the most times. To inform this selection, we also provide statistics such as the number of attempts, incorrect answers, and passes the user has had for the question. The user can then practice this specific question as they would in the game, but must assess the correctness of their answer themselves.

4.7 Statistics

To encourage users to take quizzes repeatedly and insight to their learning, various statistical measurements have been implemented. These measurements provide users with insights that can be viewed at the end of a quiz, including statistics specific to that quiz and quiz combinations, as well as on the home screen. The home screen displays game history statistics and measurements that are converted into text recommendations.

4.7.1 Game- Performance & Effort

Users can view their performance and effort on both the home screen and after completing a game. Performance is calculated by comparing the number of correct answers to incorrect ones. Effort, on the other hand, is assessed by examining the ratio of questions answered to the number of times the user has passed. These measurements are presented through pie charts.

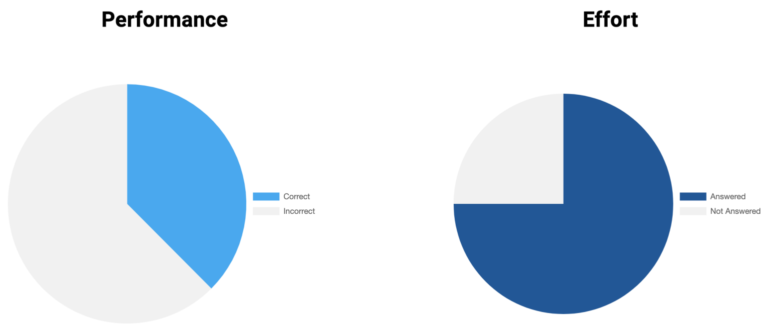


Figure 30: Performance and effort chart of a game played

Figure 30 presents a user's performance and effort for a game. Performance compares the amount of correct and incorrect answers. The effort shows the number of questions attempted compared to the total amount of questions the user has answered.

4.7.2 Forgetting Curve

The forgetting curve is employed to provide users with an approximation of the amount of information they are likely to retain in days. This estimate is represented using a line chart. If the user has previously played games that include the same quizzes, the chart will account for these repetitions, as explained in Section 2.1.2. The curve also computes the level of memory retention for the next seven days if the user takes the quizzes today, Figure 31, serving as an incentive for users to repeat the quiz. The retention rate is adjusted for each repetition as visualized in the graph.

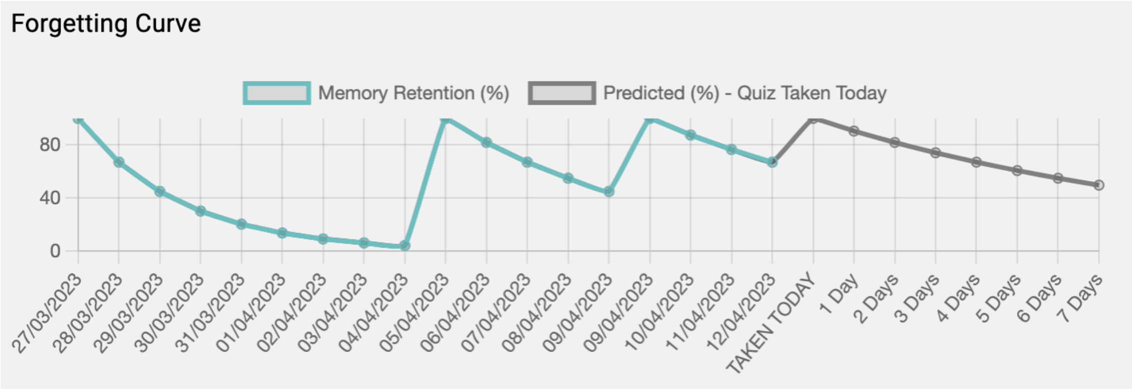


Figure 31: Forgetting curve with historic data.

The chart depicted in Figure 31 indicates that the user first took the quiz on March 27, 2023, and subsequently repeated it on April 4 and April 9, 2023. The grey line represents the estimated retention rate over the next seven days if the quiz(zes) were taken today.

4.7.3 Recommendations

For the home page we made a list of recommendations that are different for each user. These recommendations are a combination of feedback, encouragement, and other important messages that the user should consider when using this learning platform, Figure 32. The recommendations are clickable and navigate the user correspondingly. The recommendations can be categorized into incentives presented below.

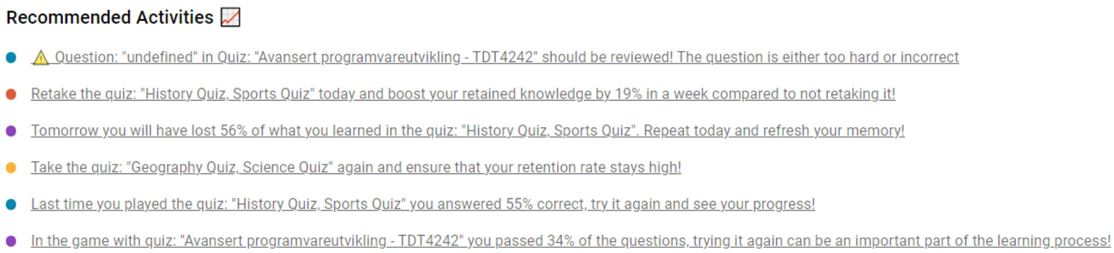


Figure 32: Recommendations displayed on the home screen

Dynamically generated - The recommendations are dynamically generated, as all the recommendations are dependent on what the user has done and how the user has performed.

Statistically based - The statistically based recommendations attempt to simplify and explain some recommendations based on the numbers we calculate by using the forgetting curve, looking at their performance, and their effort. We give the user feedback on which quiz they have the least effort on, their performance on their latest game, the quiz they will have retained the least of the material learned, and how much more their retained knowledge would be in a week from now if they retake the quiz today.

Motivational - In addition to the statistical feedbacks we implemented some recommendations that simply track a streak of quizzes or give simpler feedback where it finds the user's most taken quiz and encourages the user to retake it to ensure their retention stays high, without elaborating with what the current retention rate is estimated to be at and how it will degrade within a week.

Quiz updates - As discussed in subsubsection 4.5.1 we have a scheduled function updating the difficulty matching our modified facility index from Moodle. If one of the questions lands within the 5% index, the user will get a warning at the top of their recommendations to revise that question. By clicking this recommendation the user will be sent to the edit page for their quiz with the question at hand autofilled in the edit-box. Once the question has been edited the question will get a new identifier, meaning that the question will be treated as a new question and all future answers to this question will be handled as a first-time answer and will be open for adjustments by the scheduled function.

5 Evaluation

After following an iterative design process for implementing techniques and ideas to address the research questions a qualitative approach was employed to gain insights from participants through user tests. In total, there were conducted nine user tests that included 14 participants. In order to replicate an ideal utilization of the platform's features, we had four participants test the application over an extended period of time. The remaining participants tested the platform a single time. In this case, the main focus was to observe and gather insights into social interaction, adaptive difficulty, and support for recall in the game flow. In order to look at the effect of the proposed solution, Huskedu, the evaluation focus to answer the following questions:

- How did democratic voting of an answer work in a QA system?
- How did the different mechanisms, regarding repetition, affect participants?
- What results came from including open questions?
- How did our implementation of adaptive difficulty affect users?
- What impact does the interleaved- and blocked practice have on players?
- How does the feedback system impact participants?
- Did the implementation of game flow motivate and engage players?
- How was the social dynamic when playing quizzes?
- What improvements could be made to the current solution?

From observations and in-depth post-interviews, the insights highlight challenges, improvements, and effectiveness of mechanisms in Huskedu.

5.1 Qualitative Approach

To address the questions we followed a qualitative approach to gather detailed data on users' experiences and perceptions of Huskedu. By taking a more in-depth, exploratory approach combining interviews and observations, we were able to gain a richer understanding of how users interacted with the platform and identify specific areas for improvement that could enhance the overall usability and effectiveness of the platform. This would have been difficult to obtain through quantitative methods alone.

5.2 Conducting User Tests

During the user testing process, it was observed how users interacted with the platform and the extent to which digital and social interaction influenced the learning outcomes. The primary method of error identification involved qualitative data collection. Initially, participants were questioned about their previous experiences using learning tools. Following this, users engaged in a game that combined multiple quizzes. Throughout the gameplay, we as observers, took notes of the users' actions on the platform and their interpretations. Upon completion of the game, post-questions were posed to the participants in order to gain a more comprehensive understanding of their thoughts.

During the test phase, two main approaches were employed to conduct the user tests. The first approach involved conducting tests on different groups of users to get perspective and feedback on the platform. This approach allowed for a comprehensive evaluation of the platform's usability and effectiveness from different points of view.

The second approach, which was performed simultaneously with the first approach, involved testing a group of users every week for four consecutive weeks. This allowed observation of changes in user behavior and the impact of the platform over time. Testing users repeatedly enabled us to some extent assess the platform's potential for long-term engagement and its ability to facilitate sustained learning outcomes.

Both test scenarios started with an introduction of the game concept, followed by playing a game of combined quizzes. Afterward, the users reviewed their statistics and feedback, completed the "Daily Repetition", and created their own quizzes with both open and closed questions. The participants that tested over an extended period of time played the same quizzes for each test.

5.2.1 The Questionnaire

The questionnaire, presented in Appendix B, served as a valuable instrument for gathering systematic feedback from participants during user testing. This feedback facilitated an assessment of the platform's usability, functionality, and effectiveness in promoting learning within a social context. In the post-interview we could use this questionnaire as a starting ground for getting better insight, and query further if there were some uncertainties. In the initial user testing, we discovered that some questions could be misinterpreted and were therefore adjusted for further testing.

5.2.2 Introduction Quiz

During the initial stage of user tests, participants were asked to join a lobby and start playing the game. However, feedback from post-interviews revealed that some participants did not understand how to answer their initial questions. This was due to misunderstanding the game flow, which they understood within getting their second question. This misunderstanding gave inaccurate results to the difficulty algorithm which led to miss calculating the participant's knowledge level.

To resolve this issue, an introduction quiz was created to onboard users with five informative statements about how the game logic works. For instance, a question would be "When you click finish, there will be a democratic voting system to determine if the answer is correct". The answer would be "Was the answer correct? Discuss and click what you think is right". This approach was meant to help users focus on learning and eliminate confusion.

However, when tested in practice, participants found the informative questions to be confusing. Instead, they preferred an onboarding quiz with very easy questions such as "What is 1+1?" with the corresponding answer "2". This had the same function as the informative questions, but it was simpler. Combining this feedback with the observations of seeing the users create quizzes in previous tests, we decided to have the users create their own quizzes in the beginning instead of creating them at the end of testing. The quizzes included both open and closed questions, to get a better understanding of how they work. They were then asked to create a lobby and play against each other using their newly created quizzes. This approach allowed us to examine the ease of use in creating quizzes, which was previously done at the end of the user test. It also enabled us to see how users navigated the platform. As a result, participants got a greater understanding of how the game flow works as they have a better relation to the questions and answers. Eliminating this confusion provided us with a clearer opportunity to observe the implementation of our techniques and ideas.

5.2.3 Quizzes for User Testing

For each user test, we prepared quizzes with topics that were likely to be unfamiliar to the participants in order to test the platform's ability to improve their knowledge. This approach was more effective for participants who repeated the test over an extended period of time because it would gradually expose them to new knowledge and challenge them to learn more. With each repetition, they gained more familiarity with the topics and increase their overall knowledge. However, for participants conducting the user test only once it led to confusion and decreased learning outcomes because they did not have the same opportunity to gradually build their knowledge base.

For example, a question that was considered easy by the test creators might be perceived as difficult by the participants, resulting in a higher probability of achieving the frustration state 2.3. The ideal use case for the platform would be for users to repeat previously learned material to reinforce their learning. Therefore, it was decided to create quizzes based on topics that the participants were likely to have some prior knowledge of to maximize the effect of the difficulty algorithm. Since most of the tests were conducted with four participants in total, four questions with the estimated difficulty medium were added to the quiz. This was to ensure that each participant started on the same difficulty level.

5.3 Design for Playful Environment

As mentioned in 3.7, we wanted to create a playful environment for the process of playing quizzes. In the post-test questionnaire, we ask participants about what they felt about the design and if they had any suggestions. The feedback clearly showed us that the use of contrasts, colors, fonts, shapes and toon-based avatars was helpful in creating a playful environment and that the user felt that it made the game more fun and enjoyable. Through the incorporation of gamification elements, they were motivated to participate and remained engaged in the learning process.

Furthermore, we discovered that navigating through the website was surprisingly easy for the users. In addition to just playing the game, we let the users create quizzes, view the quizzes, and host and join games. The participants did not run into any difficulties performing these tasks and pointed out that the colorful navigation bar clearly indicated what the user must do to accomplish their desired outcome.

Overall, playful design elements and user-friendly features contributed to engaging and motivating the participants in the learning process.

5.4 Creating Quiz

Throughout the testing, participants were assigned the task of creating a quiz to evaluate the ease of navigation and quiz creation. According to post-interviews, all participants found the task easy to accomplish, but there were variations in the way they used and interpreted quiz creation.

5.4.1 Estimating Difficulty

From testing the prototypes, subsection 4.1, we received feedback that estimating a question's specific difficulty was challenging, especially if they did not know the subject very well. As we mentioned we revised some of our titles and design based on feedback from the prototypes, and for this case, in particular, we changed the title to "Estimated Difficulty", clarifying that this is just an estimate and is not absolute. We also added the helping modals where we underline how this is an estimate, and that we will automatically adjust the difficulties after the quiz has been played by a certain number of users. With these changes, we did not observe any difficulties for the users to set an estimate of the difficulty.

5.4.2 Setting a Time Constraint

We see that estimating a time constraint for questions can also be difficult. Therefore, we suggest a solution for implementing automated cloud functions to adjust these time constraints. Especially for questions that often run out of time without the users pressing finish. We found through testing that this only happens when a person is answering the question but does not reach the finish button in time.

5.4.3 Creating Open- and Closed Questions

The intended idea of adding answers to open questions was to include bullet points that the user must cover in their explanation 4.3.3. However, from the tests, it was observed variations in how users created and interpreted open questions. For example the closed question "Name three of the seven deadly sins". would be utilized as an open question, by some users, to get the possibility to list all the sins in bullet points. Even though this is not an open question by definition, it is a new way of querying questions.

The participants creating these types of questions did not use the helping modal, and was prefixed on the belief that their question was the correct use of open question. When questioning these participants we found out that most of their background associated with quizzes are that they are short and have concrete answers, such as closed questions. When asking these participants they answered that these associations are from their experience with quizzes. Some from quiz books and some from playing quizzes with friends at parties etc. Other participants were familiar with quizzes in terms of multiple choice, such as Kahoot 2.7.2 or other digital tools used within the Norwegian school system.

When asking the participants that misinterpreted the use of open and closed questions about using the helping modal we got different results. Some did not see the helping modal at all, some meant that they had no need for it as they already knew what open and closed questions are, and lastly, some opened it and felt that the explanation was too long to be bothered to read it.

On the other side, we had some users generate open questions where they used the bullet points as multiple choice, thinking that the answer would be displayed during the question. After they used the modals they changed questions to closed question and only provided the correct solution.

As mentioned in subsection 5.1, our research does not focus on this case specifically, as playing the game and answering the questionnaire already takes a lot of time. We anticipated that by playing the game they would get an understanding of how open questions work and how to use the bullet points to cover the answer topics that should be provided for an open answer.

5.5 Quiz Details

As discussed in the implementation of the view specific quiz page subsubsection 4.3.4, users can access individual quizzes to obtain an overview of specific details, including estimated completion time and average difficulty. In this section, we present our evaluation of the view specific quiz functionality based on findings from testing. We discuss unexpected observations that were uncovered during testing, as well as areas that were previously overlooked. By identifying these issues, we gain insights into how the view specific quiz page can be improved to provide a better user experience.

5.5.1 Time Estimate for Quiz

When viewing a specific quiz we include an estimate of how long the quiz takes, the current solution does not reflect the quiz's true length very well. From testing we can see due to social interactions, reflections, and administrative interaction the quizzes takes longer. A better solution would be to add an additional cloud function where we track how long the users on average spends on taking the different quizzes and change the average time for the different quizzes. Another solution would be to look at all the quizzes together and calculate the average time spent on social interactions, reflections, and administrative interactions, add this on top of the current calculations based on how many questions there are.

5.5.2 Average Difficulty

When viewing a specific quiz the users can see the average difficulty based on the questions the quiz includes. The current solution calculates this on creating the quiz and whenever the creator changes it manually, however, it does not take into consideration whenever a question is updated automatically by our cloud functions.

5.5.3 Hidden Answers

When creating the pages to view a specific quiz we opted to hide the answer behind a button click, meaning that if the users want to see the answer to a question the person must click a button to reveal the solution. This was to let users view potential quizzes to play without revealing information that could lead to advantages for future participants.

However, during an initial user test, it was observed that a quiz could be taken without playing a game. The test participant accessed the quiz screen and selected an intriguing quiz. They proceeded to read the question aloud, provide an answer, and then click on the option to reveal the correct answer.

The undiscovered new method for effectively administering a quiz and verbally stating answers seems to offer potential as a helpful strategy for reviewing information and evaluating the appropriateness of quiz material based on one's level of expertise. The person discussed the useful aspect of being able to quickly see the level of difficulty for each question. Additionally, they pointed out that this feature could be easily used to quiz oneself or others in preparation for an upcoming test.

5.6 Game Flow

In this subsection, we discuss different aspects of game flow, including scorekeeping, time constraints, repetition phase, and some edge cases. We discuss different scoring systems and their impact on the competitiveness of the game. We also consider the issue of uneven teams and how to adjust the point system to ensure fairness.

Furthermore, we suggest solutions to address time constraints, repetition phases, and edge cases.

5.6.1 Scoreboard & Delegation of Points

As outlined in subsubsection 4.4.8, teams can receive both positive and negative points in the game. Based on observations and post-interview feedback from the user test participants, the scoring system has the potential to increase or decrease competitiveness. The allocation of negative and positive points widens the gap between teams, making it more challenging for a team that is behind to make a comeback. Such a scoring system is ideal for scenarios where teams have similar levels of knowledge. However, in cases where there are significant differences in knowledge, it would be more suitable to assign zero points if an answer is incorrect or passed.

One insight gained from the user test is that different point allocations could be applied to various aspects of the game. For example, in the repetition phase, answering incorrectly incurs a penalty of -10 points due to the fact that the answer has been seen before. In all other phases, teams only receive zero points for incorrect or passed answers. This penalty system in the repetition phase could enhance user concentration since they are penalized more severely for incorrect responses.

Additionally, a participant suggested a novel idea for allocating points in open questions. As currently implemented, users' responses must include some or all of the bullet points listed in the open answer. The other team members can then checkmark the bullet points the player answered. The allocation of points would for example correspond to the number of bullet points.

When creating the platform's game flow we added the possibility for the teams to be uneven. This felt natural as we do not want to limit the platform to only groups of even numbers. As our focus was on letting each player receive the same amount of questions from each category, playing with uneven numbers leads to unfair game play. The team with less players receives less questions than the other team, therefore, has lower maximum points achievable than the other team. But with the inclusion of subtracting points, the team with less players also has lower maximum negative points achievable, which potentially evens the score out.

When discussing this mechanic with the participants that played with uneven teams, some felt that when focusing on the learning aspect it is important that each player gets the same amount of questions, but focusing on the score, it felt unfair for the users with less players.

A solution would be to adjust the points based on the number of players in each team. Since the difference between players will always be one, we can subtract ten points for the team with fewer players from the default points allotted, which is divided by the number of players on the smaller team. For instance, if one team has two players and the other has three, the team with three players will keep getting 10 points, while the team with two players will receive $10 + 10/2 = 15$ points. This ensures that both teams have equal potential for minimum and maximum scores, and all players get the same number of questions, regardless of mini-games.

5.6.2 Open Questions

As noted in subsection 2.7, many learning platforms available provide tools for optimized learning outcomes. However, these are most commonly combined with closed questions. In the testing phase of Huskedu, the participants found the option to answer open questions to be an unfamiliar approach. However, post-interviews showed that open questions are often used in oral and written tests and are especially popular for test preparations.

Feedback from users highlighted the benefits of incorporating open questions in a quiz platform. They pointed out that open questions support personalized responses, as they allow them to provide unique explanations rather than limiting them to a specific answer. This prompts self-explanation, subsubsection 2.2.2, and can lead to better learning outcomes. Additionally, users noted that open questions require a more comprehensive understanding and knowledge of the material at hand, which enables them to be tested on questions that require greater understanding.

Furthermore, post-interviews indicated that incorporating open questions can prompt discussion and collaboration among users, as opposed to closed questions. This leads to a more engaging and interactive learning experience where users can share different perspectives and learn from one another. As a result, new learning material can be illuminated, expanding the user's knowledge base.

5.6.3 Voting

According to feedback received during post-interviews, the democratic voting system used to determine the correctness of answers was found to be a feature that generated engagement and sparked discussions among participants, especially for open questions. Our observations showed that players sometimes had to defend their answers by presenting arguments. This requirement led to a situation where the individuals concerned had to provide reasoning and evidence to support their claims. As a result, any gaps in their understanding and weaknesses in their arguments could be identified. This process aligns with the principles of self-explanation theory, subsubsection 2.2.2, which suggests that it is a powerful approach to optimizing learning outcomes.

Since all answers were determined to be correct or incorrect, through voting, all players are required to pay attention when a player elicited their answer in order to categorize the answer as correct or incorrect. This led to several beneficial learning outcomes, including active engagement, accountability, and peer learning. All participants revealed in the post-test questions that they felt they paid attention when others are answering, which promoted active engagement with the material being discussed. The process of voting creates accountability for each player's answer, which encourages them to take their responses more seriously and put more effort into providing accurate and well-reasoned answers.

Furthermore, with the combination of open questions 5.6.2, the open questions sometimes led to deeper discussions, where participants discussed whether the answer provided by a user was significant enough, and if they used other terms than the solution included. One participant explained his relation to open questions as this:

> “When providing an answer for open questions we often include some correct and some incorrect information, and can be challenging to give a concrete answer”.

The participant further explained that the process of figuring this out was an overall positive experience. Situations like this prompted users to go even deeper into questions, taking advantage of search engines to justify their argument or simply discussing internally and elaborating their explanations and the terms they used.

Interestingly, we observed that participants always argued in favor of the questioned participant, regardless of whether they were teammates or not. This created a friendly environment where participants helped each other to better understand the material. Moreover, discussions often went so far as to accidentally cover future questions from the same quiz. This required participants to use retrieval practice subsubsection 2.2.2 to recall information recently discussed from memory when they got a question about what they recently discussed. This helped them answer correctly and deepen their understanding of the material.

In conclusion, through the process of voting and discussion, players have the opportunity to learn from one another’s perspectives and insights, which can help to broaden their understanding of the subject matter and expose them to different ways of thinking about the material. Furthermore, The inclusion of a democratic voting system and open questions led to unexpected and valuable discussions, elaborations, and deeper understanding of the topic than the quiz provided. We found that these two mechanisms complemented each other well and had a good dynamic that enhanced learning outcomes.

5.6.4 Time Constraint

Time constraints contributed to an increase in engagement by adding a sense of urgency encouraging players to actively participate and stay focused on the game. When a player would give their answer they were forced to give concise and clear answers including the most important aspects of answering the question. This prompted players’ critical thinking and problem-solving skills under pressure. Overall, the constraints played a part in setting the pace of the game. Time limits helped with a lively atmosphere and prevented the game from becoming slow-paced. However, as the question’s time constraints were an estimate from its creator sometimes it did not suit the question as it was either too- high or low.

An additional observation from implementing time constraints is its effectiveness in facilitating the use of open-ended questions. Interactions with participants revealed that determining the appropriate level of detail for an open-ended question can be difficult, but the time constraint proved to be a helpful guideline. We found that the time limit for an open-ended question provides an indication of the expected level of detail. Moreover, when a question had a short time limit but required a detailed response, the subsequent discussions after the question often took this factor into account when evaluating the accuracy of the participants' answers.

5.6.5 Repetition Phase

In our post-interviews, it became evident that the repetition phase of the game was one of the most attractive aspects of Huskedu. Users consistently liked the game's approach of revisiting previous questions, with one participant stating,

> "Very nice! It is fun to answer questions that both I and others have answered incorrectly on, and rewards players to pay attention to each other's answers".

The repetition phase's approach of first introducing questions that players have answered incorrectly before focusing on those others players had answered correctly was viewed as highly effective for optimizing learning outcomes and engagement. By revisiting questions that were answered incorrectly in the past, learners were able to reflect on their mistakes and gain a better understanding of their errors and gaps. When the participants were asked questions that others had answered correctly on, it became beneficial in the social context due to each other's previously elicited insights and knowledge of the subject matter.

We observed that the repetition phase placed at the end of the game was highly effective as all participants had a good understanding of the game's flow and logic by this point, making it easy to participate. Additionally, we found that participants who had tested the platform over an extended period of time paid extra attention to the questions on the following user tests as they knew a repetition phase would follow at the end of the game. The participants also noted that the competitive social play aspect of the game encouraged them to concentrate and engage with all questions, especially when they knew the repetition phase would require them to recall previous answers to win. This combination of competition and social play, along with the repetition phase, was highly effective in keeping learners engaged and motivated.

5.6.6 Game Administrator

After the first two tests with participants we changed from being the administrators ourselves to letting participants be the administrator of the game. By being administrator one controls the pace of the game, which had an impact on the social dynamic of the game. Most participants who were administrators of the game became more engaged and vocal. They repeatedly asked if the users are ready for the next step, or announced that they will now start the next part. When asking the participants what they felt about the role, they clarified that they felt that this role was engaging and was easy to fulfill.

We observed one scenario where the administrator moved to the next part too early. The other participants felt that they were not done discussing the solution, but quickly let the administrator know that they moved on too fast and wanted more time to discuss. However, at this point the game has moved on and the discussion of that solution did not resurface.

In the earlier stages of developing the solution of how the game should be controlled we discussed an automated timer. The problem with an automated timer is that it sets a time constraint to discussions and counteract the idea of having discussions.

A suggestion that was brought up during the post-interview by one of the participants was to use the same concept of democratic voting used for determining if the users answers correct or not, for the administrative role. This idea has good potential as it keeps all users engaged throughout the game. However, a concern is that it might be a lot to focus on and that it could lead to confusion in regards to why there is another voting system. Another concern is that if one of the players regularly forgets about having to vote to move on to the next step could lead to frustration among the other participants.

Through our observations we can tell that with the administrative role, there is a potential to further engage users. However, administrators that do not care for discussions could potentially be a problem. Through our observations, this was a rare case and was adjusted quickly. Therefore, we feel that the current solution is optimal as it impacts the social dynamic positively, but if this became an issue in the future, a solution could be to let users choose between having an administrator or using the automated timer with the addition of buttons for pause and skip. Where the pause button can be used to discuss further, or a skip button that the majority of users must click to go to the next step.

5.6.7 Edge Cases

Tie votes - During testing, a recurring issue among subjects was the allocation of points when the number of correct and incorrect votes for an answer was equal. In the current system, a tie results in the allocation of +10 points, as stated in subsubsection 4.4.8, to the team of the person who provided the answer. However, some participants noted that this could be exploited by the answering team to their advantage. Consider an example of a game featuring four players divided into two teams with two members each. Suppose that a participant from the Yellow team submits an answer, confident of its accuracy, but both members of the Purple team mark it as incorrect. In such a situation, the final member of the Yellow team possesses the authority to determine whether their team should be rewarded with +10 or penalized with -10 points. Given that the person who gave the answer belongs to the Yellow team, they could potentially cast their vote in favor of their team, thereby securing a +10 point allocation.

To prevent potential biases in point allocation, one possible solution is to restrict the player who answered the question from voting on their own response. However, if the game has an odd number of players, this could result in an equal number of votes, resulting in a tie. In such cases, a more equitable option would be to assign zero points for the tied responses, thus avoiding any undue advantage or disadvantage to either team.

No participation in mini games - During the user tests, a potential issue was identified in the mini-game phase where a difficult question was presented. As currently implemented, at least one participant must complete the mini-game in order to proceed to the next question. However, if none of the users are willing to attempt the game, the quiz cannot progress. This challenge was observed during the tests and highlighted the need for one participant to take the lead and receive a penalty of -10 points in order to move the game forward.

No mini-game scenario - With the current algorithm subsubsection 4.4.1 there exist two specific scenarios where the minigame phase will not be included, this is because the algorithm prioritizes that each participant gets the same amount of questions from each quiz. The conditions for this scenario to happen is whenever the number of questions is the exact same as the number of participants, or exactly double the number of participants. This was intended, as our main focus is on the quiz itself and not the mini-games.

The removal of mini-games did however create a shift in the social dynamic of the game, losing some of the engagement that the mini-games created, this is further discussed in subsubsection 5.7.2. Even though this was intended, we did not anticipate it happening when combining quizzes, as the chances for this are very low. A solution to the scenario where both quizzes are exactly two times the length of the number of participants, we could take half of the questions from one of the quizzes and make this into a mini-game, this way everyone would still have the same number of questions from each quiz. A weakness with this solution is that if some participants are better at one quiz than the other, it could lead to a disadvantage to that team in particular, as the questions are up for grabs for every participant.

Repetition question delegation Choosing repetition questions could be influenced by other players performance. As the repetition questions are turn-based and chosen in the order introduced in subsubsection 4.4.7. This means that if user A answers all correct and user B answers 2 incorrectly, user A will get one of user B's questions and user B will only get one of the questions they answered incorrectly in their repetition questions. This could both be positive and negative, as user B will be witnessing a potentially more knowledgeable player give an answer on something him/her-self answered incorrectly, but also robs the person of the opportunity to attempt to solve the question themselves.

5.6.8 Race Conditions

Real-time updates are a critical component of ensuring equal participation amongst all users in the quiz. The initial user test, involving two fellow students, exposed challenges associated with race conditions that arose during the process of joining a lobby and casting votes in the game. When the users attempted to join the lobby, their simultaneous actions led to a race condition in the database, resulting in only one user being registered in the lobby. During subsequent gameplay involving two quizzes, multiple instances occurred where only certain votes were recorded. This, too, was attributed to race conditions that hindered updating answers with votes.

Following feedback from the initial test participants, who identified this as a relatively significant and frustrating issue with the platform, it became necessary to address these race conditions. Ultimately, Firebase's transaction methods were leveraged to resolve these issues.

5.7 Social Dynamic

From the user test, we explore the impact of different social environments and features, such as friendly compared to competitive atmospheres, mini-games, and repeated gameplay. The findings suggest that the social dynamics of the game are affected by the level of competitiveness and the presence of mini-games, which promote engagement and fun. Additionally, repeated exposure to the game seems to enhance participants' understanding of the game's mechanics and rules, leading to more productive discussions and interactions.

5.7.1 Testing in Friendly Environments

The user test post-interviews revealed that different social settings may have a substantial impact on how players are affected by one another. Specifically, when played in a school context where players are randomly assigned, the game demands a higher level of performance than when played with friends or family. The majority of user tests were conducted with two or more participants who already knew each other before testing. This meant that most of the test scenarios were conducted in a friendly and comfortable atmosphere, where participants felt at ease providing incorrect answers and engaging in discussions. On the other side, a tendency to engage in humor was frequently present than anticipated in these situations.

When discussing this with the concerned participants, they pointed out that in a more formal setting, such as an academic environment, a more serious tone will arise. This tone might further promote focus on the learning outcome rather engage in humor.

5.7.2 With & Without Mini-games

A minor subset of the user tests involved omitting the mini-game component of the game exploring the effects of mini-games and competition on learning outcomes in an interactive learning QA system. The purpose of this was to evaluate the impact of breaks and competition on learning outcomes for the participants.

> "Fun and engaging! Sparked my competitive side"

According to the post-interviews of the participants, the mini-games played a crucial role in enhancing participation and engagement on the platform. The mini-games fostered a greater focus on the score and promoted a competitive and fun environment within teams. On the other hand, participants who did not play the mini-games expressed that they did not find them necessary and even considered them to be a disruptive feature. However, they also noted that the inclusion of mini-games might be appealing in friendly settings.

5.7.3 Repeated Games Over an Extended Period of Time

Participants who played games repeatedly over an extended period of time displayed a more visible social dynamic compared to single-time participants. This is likely due to their clear understanding of the game logic, making the flow of the game flawless. This pre-existing knowledge led to more discussion among participants. Having a better understanding of the game's mechanics and rules may have enabled them to engage in deeper discussions, which likely increased the social dynamic of the game. Their enhanced confidence in participating in discussions and providing input may have encouraged other participants to engage in the conversation, leading to more active and productive discussions. Additionally, repeated exposure to the game may have built a stronger relationship among participants, further enhancing social interactions and discussions.

5.7.4 Facilitate Collaboration Within Teams

During the user testing of gameplay, it became apparent that discussion among participants was prompted by the accuracy of answers. Additionally, it was observed that the teams were not working collaboratively, except for contributing points to the overall score. To encourage both team collaboration and vocal discussion, it is suggested that a feature be incorporated allowing teams to be delegated questions rather than a single player. The difficulty level of the delegated question could be intentionally challenging to prompt discussion, or calculated based on the team's previous average performance.

5.8 Blocked- & Interleaved Practice

The implementation of blocked and interleaved practice had a notable impact on various aspects during gameplay. Through user testing, we observed different advantages and flaws associated with each practice approach. This highlights the importance of considering diverse practice methods to optimize user motivation and information retention. By analyzing user feedback and performance, we gained valuable insights into the benefits of incorporating both blocked and interleaved practice.

5.8.1 Embedding Multiple Quizzes

Without the possibility of embedding multiple quizzes, there would not have been possible to utilize blocked- and interleaved practice. Most of the user test participants found the possibility of embedding multiple quizzes to be an attractive feature. Furthermore, they felt that this would allow them to practice multiple topics or sub-topics at the same time.

5.8.2 Comparison of the Practices

The majority of the tests were conducted using a blocked practice approach, primarily due to the topic's differences. Furthermore, participants found it easier to answer when they were asked questions on the same subject. This could be attributed to the fact that similar questions were asked previously, or that some questions triggered the retrieval of additional information related to the subject matter. Post-interviews revealed that repeating a quiz a week later using blocked practice enabled the participants to recall the order in which questions were asked to some extent. However, interleaved practice eliminated this assistance. Observations from the interleaved practice indicated that participants felt they learned better since they were required to change and retrieve information on various topics. Furthermore, differentiating topics allowed for the identification of similarities between them such as explaining the structure of a government.

5.8.3 Impact on Game Flow

Comparing the two practices integrated in the game flow, as shown in Figure 33, we noticed a difference in engagement and motivation. In the blocked practice, mini-games were after each category, allowing users to get small breaks and remain engaged. This approach encouraged user participation and sparked competition amongst users.

However, in the interleaved practice, all the mini-games were gathered at the end of the game. This approach had the opposite effect on user engagement, especially when the game contained many questions because the different phases came in batches instead of being intertwined. The same goes for the introduction phase, as users were repeatedly presented with questions. Additionally, when users were prompted to play the mini-games, they had to play many games in a row, which ultimately resulted in boredom. To address this issue, mini-games could be spread out during the introduction phase to maintain user engagement and motivation throughout the game in an interleaved practice. A revised suggestion for the interleaved practice game flow is illustrated in Figure 33.

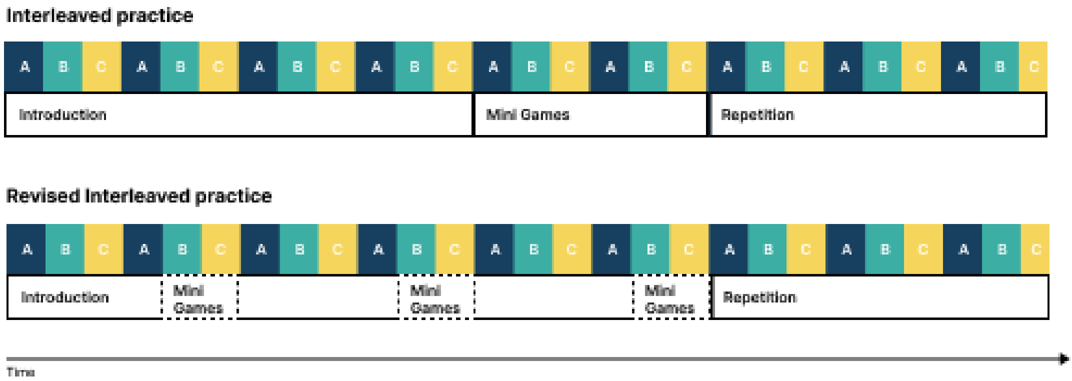


Figure 33: Comparing the original interleaved practice, in a game, with the revised version that has spread out mini-games.

Distributing the mini-games throughout the introduction phase can be advantageous for learning outcomes as it would add variety to the introduction and maintain user engagement. This engagement could lead to improved retention of the information being presented. In contrast, when mini-games are clustered at the end of the game, users may become bored or disengaged, which could lead to decreased retention of information and reduced motivation to participate in future learning activities.

5.8.4 Impact on Motivation

During specific user tests, participants had varying levels of knowledge about the topics covered in the quizzes. It was observed that in blocked practice users were repeatedly asked questions about the topic they knew the least about leading to a decrease in motivation. Conversely, when participants were asked questions about the topics they were more familiar with, their motivation increased. This dynamic also affected the social engagement of some users, with some being more focused on their individual performance while others were more concerned about the group's performance as a whole.

In contrast, when users engaged in interleaved practice where every other question was from a different topic, their motivation remained stable. This was because they received an equal number of questions on their preferred and non-preferred topics, which helped balance their motivation.

5.9 Adaptive Difficulty

The user testing showed that multiple participants played games with various difficulties. The findings show that the delegation of appropriate difficult questions was impacted by various factors that would be optimal to incorporate in the algorithm to improve accuracy. In the following subsection, we will present the challenges with the currently implemented algorithm and look at possible improvements.

5.9.1 Estimated Difficulty

All the questions have an estimated difficulty, leading to a sub-optimal adaptive difficulty. With the current solution, all question difficulties are estimated by the creator of the quiz. This works well in some cases, but it is dependent on what perspective the creator has. Is it created for a group of friends that are studying the same subject, or will the quiz be played by people that has no relation to the subject. The adaptive difficulty algorithm takes use of the estimated difficulty when delegating questions, therefore, it can delegate correct in theory but as the estimated difficulty might be wrong it can be presented as an incorrect delegation. This is something the cloud function subsubsection 4.5.1 would be able to adjust after enough participants has played the quiz.

5.9.2 Automated Difficulty Adjustment

Since the scheduled cloud function for automatically adjusting difficulty is based on unique first-time attempts from users, it would require more unique participants to get a justified result. Due to the limited amount of test participants, this was not possible to achieve. As a result, we simulated the cloud function and verified both the result compared to the facility index, and the recommendation for questions that lands within the 5% index, subsubsection 4.7.3.

The simulation was done by running a quiz 20 times, answering six different questions correct/incorrect with a distribution reflecting the different indexes in the facility index subsubsection 2.5.2.

5.9.3 Open- Compared to Closed Question

In general, it was observed that open questions tended to be more challenging than closed questions. For instance, if a closed and an open question were estimated to be equally difficult, the open question was still more challenging since the user had to remember and articulate the answer in their own words, covering all sub-answers to get it right. Conversely, for a closed question, the user only needed to recall a specific answer. To address this, the delegation algorithm could be adapted to alternate between open- and closed questions. This approach would ensure that all users get to answer both open- and closed questions.

As open questions seemed to be more difficult compared to closed questions we want to highlight that including both these types of questions enabled users to show both specific and detailed knowledge. From observations closed questions were representative of measuring specific knowledge, as they usually had a clear and concise answer, being useful in evaluating a user's ability to recall facts, definitions, and concepts. On the contrary, open questions required users to provide more detailed and nuanced answers, which demonstrated their ability to apply knowledge, analyze information, and think critically.

As discussed in subsubsection 5.6.1, it may be useful to consider the possibility to select bullet points that the person concerned included in their answer. This feature could be incorporated into the difficulty algorithm in the repetition process such that answers that covered fewer bullet points are more likely to be repeated than those incorrect answers that covered more.

5.9.4 Question Difficulty Distribution

The distribution of questions with various difficulties in a quiz is critical for the accuracy of the algorithm. If a quiz is composed primarily of questions with a specific difficulty level, such as "Hard", the impact of the algorithm will be minimal. Quizzes that are mostly comprised of hard or very hard questions tend to result in a state of frustration, while quizzes that have mostly easy or very easy questions lead to boredom.

Quizzes should have a variety of difficulty levels, ideally, this should be a normalized distribution. This balance of difficulty levels will enhance the overall experience of the quiz by keeping the players engaged and motivated, while increasing the chance of users entering the flow state.

5.9.5 Utilizing Historic Data

The current algorithm for determining the optimal difficulty level of a question only considers data from the current game. As a result, participants who test the application over an extended period of time may encounter questions that they have answered before. This is because the algorithm only compares data from the current game to select questions of appropriate difficulty. To address this issue, an optimization could be made to leverage previous historical data. By considering a user's previous performance on certain questions, the algorithm can suggest new unanswered questions of appropriate difficulty.

One approach could be to create a personalized difficulty model for each user that maps questions to difficulty levels based on their past performance. For example, if a user has answered a difficult question, estimated as hard, correctly multiple times the model could map that question to a medium difficulty level for the user concerned. This approach helps to ensure that the algorithm selects questions that can help in guiding the users toward a state of flow based on their knowledge level.

5.9.6 Equity

In user tests where the difference in knowledge level related to the subject of the quiz was high, we observed that the algorithm lead to equity amongst participants. As the algorithm adapts individually, the participants with lower levels get easier questions, further increasing the chance of getting a correct answer. Furthermore, as they get easier questions throughout the game, they also get easier questions in the repetition phase. As a result, the score balances out despite the differences, keeping participants engaged and motivated.

5.9.7 Alternative Change of the Difficulty Algorithm

One possible solution for increasing the accuracy of the difficulty algorithm could be to ask users for their subjective level of familiarity with the given topics. Prior to starting each game, users could be prompted to answer a question such as "How confident do you feel about your knowledge of Human Anatomy?". Using a Five Likert scale ranging from very little to very much, users could provide their subjective assessment. This information could then be mapped onto the scale for question difficulty within the quiz. For example, if a user reports having very little knowledge of the subject, their initial questions would be categorized as very easy.

5.10 Statistics & Recommendations

Feedback and recommendations were designed to help users improve their learning outcomes by tailoring their educational experience to their individual needs and preferences. However, post-interviews illuminated challenges in understanding some of the feedback and recommendations. This section explores both the issues that arose and the usability of statistics and recommendations provided by the platform.

5.10.1 Forgetting Curve

The feedback provided by the forgetting curve was generally well-received by users. We observed that even though the concept of a forgetting curve was unfamiliar, they found it interesting and insightful. However, there were issues with the presentation of the "Taken today" point that led to confusion among users.

The "Taken today" point, which can be viewed on the x-axis in Figure 31, was included to motivate users to retake quizzes and improve their knowledge retention. The group of participants that tested over an extended period of time, found the forgetting curve with the prediction of retaking the quiz today interesting and felt that it did in fact motivate them further to retake the quiz. One participant said

> "Excited to see the graph the next time I play".

For the users that only participated in the user testing once, we noticed that some confusion could arise. The reason for this is that their knowledge level had not yet degraded, so connecting a 100% knowledge level with another 100% knowledge level and a prediction of retention rate in the future was found to be unnecessary and misleading. To address this issue, we recommend restricting the use of the "Take today" point to instances where the user's knowledge has dropped substantially. This way, the point can effectively illustrate how users can regain their knowledge level and reduce the retention rate properly.

In conclusion, our evaluation highlights the strengths and weaknesses of the forgetting curve and the "Taken today" point. While the inclusion of the "Taken today" point is useful for motivating users to retake quizzes and improve their knowledge level and retention, it can be confusing when their knowledge level has not yet been degraded. By restricting its use, we can better illustrate how users can regain their knowledge level and reduce the retention rate properly.

5.10.2 Recommended Activities

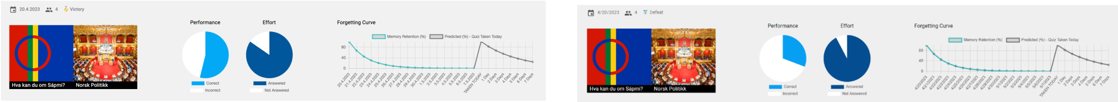
According to user test participants, the recommended activities on the platform were an engaging feature that sparked an internal motivation to improve. However, an issue among some participants was that they included an excessive amount of text and information, which made them difficult to understand. One participant said that it was too long and did not read them. While others found them interesting and enjoyed the fact that it was possible to click on a recommendation to start the process of completing that recommendation.

When discussing this further with the participants that felt it was too much information, we discussed the possibility of having a more visual representation of the recommendations. These recommendations do not go into details of why they should retake them, but display some statistics that we base the recommendations on. Participants found the thought of this solution more manageable and less overwhelming.

Despite the feedback about the amount of information displayed, the current version of the recommended activities feature fulfilled its purpose of increasing motivation and encouraging users to retake quizzes. While it did not work for absolutely all participants, it had a positive effect on most, providing a good foundation for further improvements.

5.10.3 Game History

The platform enabled its users to monitor their progress by assessing their performance as they completed games, a particularly advantageous aspect for those who tested over a longer duration. This function acted as a motivational tool, as users could compare their current game results with their previous performance and effort, often noticing an improvement that increased their commitment to the platform. For instance, a pair of users had a performance of 53% and 31%, and an effort of 76% and 91% Figure 34, the first time they attempted playing the quizzes. However, in their final test over the extended period, they achieved 100% in both performance and effort, meaning that they did not pass a single question and answered all questions correctly. Before and after each game, they compared their prior feedback, striving to surpass their previous game statistics. In terms of overall usability, every user test showed that individuals had no trouble understanding the game history feature.



(a) Player A: 53% in performance and 76% in effort

(b) Player B: 31% in performance and 91% in effort

Figure 34: Initial user test stats of those that repeated it over an extended period of time

5.10.4 Requested Features

When asking participants if there was any information they wanted more insights into, there were two topics in particular that stood out.

Questions answered - During the user testing, participants requested a more detailed overview of the questions they had answered and whether they answered correctly or not. The current solution shows how many questions they got right and how many questions they attempted to answer, without providing any insight into which specific questions were answered correctly or incorrectly.

Initially, we believed that users would navigate to the quiz page and review all the questions, regardless of whether they had answered them correctly or not. This way the users would not only look at the questions they received, but the entire quiz. However, through our observations, we found that this was not the case.

Providing users with the ability to view the specific questions they answered and whether they answered correctly or not would increase the amount of repetition, which aligns with our goal of enforcing recall from memory. Our initial assumption that users would review all questions was found to be incorrect through our observations during testing.

In conclusion, the request for a more detailed overview of questions and accuracy could lead to increased repetition and a better user experience, as it provides users with specific areas to focus on when reviewing their quiz results.

Repetition questions - Some participants requested a more indepth view of the repetition questions in particular, especially how they performed on questions that they previously answered incorrect on, but also questions they got in the repetition phase but did not receive earlier in the game.

An implementation of this has the potential to emphasize several aspects of the platform. In cases where the users answers incorrect, it can emphasize exactly how fast we can forget information presented to us. In cases where it is correct, one can get motivated and feel a sense of accomplishment. Also when there is a mix of both incorrect and correct, the users have the potential to identify specific questions that they struggle with and questions they master.

5.10.5 Daily Repetition

The mechanism of adding a daily repetition question, along with a statistical overview of the corresponding question, had a greater impact than expected. Initially, this mechanism was meant as an addition to emphasize the repetition parts of the platform. We did not expect that adding a daily question to repeat to have a great amount of feedback.

> "When I take the daily repetition question, it triggers a connection to the other questions in the quiz."

From the feedback regarding the daily repetition question, we see that a small but comprehensive mechanism like this can have a positive impact on their overall experience. Our concern was that the algorithm chooses what we believe to be one of the most challenging questions for them and that this could discourage some users. It turned out to be the opposite, all participants felt that completing the daily repetition was fun and motivating, and some said that it was a quick way to refresh and maintain knowledge. It turned out that many users were familiar with this concept from other platforms and applications, where they have a daily task to finish.

Despite our initial concern that the algorithm might choose one of the most challenging questions, users felt encouraged to complete the daily repetition. This mechanism gave users an incentive to revisit the platform and retake quizzes, potentially triggering them to repeat the quiz. Overall, adding a daily repetition question has proven to be an effective way to emphasize the importance of repetition and enhance the user experience.

6 Conclusion & Further Work

In this research, we utilized various techniques and concepts to address the research questions outlined in section 1. To accomplish this, we created a learning platform called Huskedu, which is based on QA learning. The development of the platform followed an iterative design process and underwent qualitative evaluation through user testing. These tests helped us identify the strengths and weaknesses of the techniques and mechanisms employed. Based on these results, the research questions will be answered to the best of our abilities. However, it is important to acknowledge the limitations of our work and highlight areas that require further investigation and improvement.

6.1 Conclusion

In conclusion, this research proposes a solution, Huskedu, to increase learning outcomes through repetition and social competitive play. The mechanisms in the developed platform were evaluated qualitatively through user tests, revealing strengths and weaknesses. Our implementation of an interactive learning QA system, achieved a balance between vocal social interaction and computer-based learning, all within the framework of a competitive game.

Based on research question one (RQ1), vocal social interaction and computer-based learning were combined through the requirement of answering questions vocally with a democratic voting system, an administrator choosing the pace of the game, and the inclusion of open questions. The effect of having open questions and a democratic voting system lead to more extensive discussions, in contrast to a closed question that entails a static answer. Moreover, the combination of democratic voting and the administrator controlling the pace of the game made room for discussions where players would illuminate gaps and errors and extend their knowledge by sharing different perspectives and learning from one another. In some cases, the discussions were so extensive that the participants acquired more knowledge than the solution had to offer.

From research question two (RQ2), question difficulty and the order of questions were utilized in the game as a way to optimize learning outcomes. A difficulty algorithm was created to guide users toward a flow state, engaging and motivating players to learn. The algorithm compares users' previous performance with the remaining questions and adjusts the level of difficulty based on the estimated difficulty of each question. This assisted in appropriately challenging users on an individual level. The question order is impacted by the difficulty algorithm, participants, selected practice, and game phases. The order of questions is dynamically changed throughout the game, always searching for the most suitable question for each participant. By utilizing these factors we see that the participants gain a better chance of engaging in the content regardless of knowledge level. Furthermore, combining this with equal distribution of questions, we see that this equity and equality improves the engagement, hence improving the social dynamic and learning of the platform.

Finally, in research question three (RQ3), support for recall and memory retention was implemented by emphasizing repetition and statistical measurements. At the end of each game, a repetition phase and algorithm were responsible to query users with the best possible questions that would maximize the learning outcome for the player given a player's previous response. Participants felt that this phase was a fun and engaging way to prompt recall from memory. After completing a game, the platform displays several statistical analyses to the user, encouraging them to strive for better performance and repeat the learning material. Among these measurements a forgetting curve is displayed and made users aware of their loss in memory retention, showing when it would be wise to repeat quizzes. The personalized recommendations and daily repetition gave insights into smart learning activities for coping with the challenges of memory loss. Overall, the use of repetition and statistical measurements resulted in letting users get a hands-on experience illuminating both strengths and weaknesses of memory retention and prompt recall from memory.

This research demonstrates how repetition and social competitive play can enhance learning outcomes. Huskedu, with its combination of vocal social interaction, computer-based learning, and competitive game elements, offers an innovative and effective approach to learning and discussing material. The algorithm guiding users towards a flow state, the utilization of question order, and the inclusion of repetition phases all contribute to an improved learning experience that supports recall and enhances memory retention. The statistical analyses provided motivates users to strive for better performance and reinforce the retention of acquired knowledge through repeated engagement with the learning material.

6.2 Limitations of the study

A limitation of the study was that we only conducted user tests on university students. Having feedback from a specific group of individuals lacks diversity and poses a threat to the validity of the research. Optimally we would also have tested Huskedu's techniques and mechanisms on learners from middle- and high school. This is due to their structure where the learners have frequent tests covering several distinct and indistinct topics within subjects throughout the semester. This school structure aligns with the repetition and feedback techniques of Huskedu, and offers the opportunity to see correlations between topics and subjects. Furthermore, due to the fact that students at these grades do not choose the topics themselves, it can become tedious to learn about them. In such cases, it would be interesting to see if the competitive game and social learning aspect of Huskedu would increase motivation and engagement for learning.

Another challenge we encountered during the testing phase was the fact that our platform, Huskedu, is designed for multiplayer and requires a minimum of two players to play a game. Even though playing with two players is possible, Huskedu is better suited as a game of at least four participants. As mentioned in subsection 5.2, it was difficult to find multiple participants that could test at the same time. Despite the challenges, we were able to gather valuable feedback from participants, which helped us improve the platform's usability and effectiveness and evaluate the different ideas and techniques. However, it would be optimal to conduct more tests and include more backgrounds to get a more profound result.

From the limited number of participants, we were unable to test the scheduled cloud function designed to adjust question difficulty, as highlighted in subsubsection 5.9.2. We believe that the effect of this function would have significantly contributed to accurately determining the difficulty level of questions, thereby enhancing the effectiveness of the adaptive difficulty algorithm in the game. This would have enabled us to further evaluate the performance of the algorithm and make necessary adjustments if required.

6.3 Further work

As formerly outlined, the primary constraint affecting the reliability of the findings has been the insufficient diversity in the demographic composition of the user tests and the overall number of participants. A more comprehensive testing phase, including middle- and high-school learners, would have offered a stronger basis for conclusions and uncovering novel discoveries.

Exploring the outcomes of implementing tests in real school classes for an extended duration holds a valuable interest. By collaborating with a teacher, we can develop quizzes that are well-suited for both blocked- and interleaved practice, while also obtaining valuable feedback from a teacher's perspective. Moreover, we believe that the platform is better suited for testing over extended periods of time due to its focus on repetition. Having a class that runs interleaved quizzes from two subjects over an extended period of time could bring significant insight into the usefulness of our techniques. Additionally, a teacher often has greater insights into the existing social dynamic of the class and individual knowledge level. This enables them to see how Huskedu affects the social dynamic and how it plays a role in their learning progress.

Before conducting the additional tests, it would be ideal to include some of the potential improvements discovered and outlined in the evaluation, section 5. These improvements include facilitating collaboration in teams, automating adjustments to time constraints, modifying point delegation, including statistical feedback from the repetition phase, and changing the placement of mini-games in interleaved practice. These suggestions have the potential to increase fairness, motivation, and learning outcomes.

Bibliography

[1] Anki. *About Anki*. URL: https://apps.ankiweb.net/.

[2] Punyawee Anunpattana et al. *Capturing potential impact of challenge-based gamification on gamified quizzing in the classroom*. URL: https://www.sciencedirect.com/science/article/pii/S2405844021027407#br0370.

[3] Ekaterina Barcellos and Galdenoro Botura Junior. 'Design Process: Metodologia de Interação entre Design e Engenharia como Vetor de Inovação'. In: (Dec. 2016), pp. 1307–1317. DOI: 10.5151/despro-ped2016-0111.

[4] Kiran Bisra et al. 'Inducing Self-Explanation: a Meta-Analysis'. In: 23 (2018), pp. 387–411.

[5] Paulo F. Carvalho and Robert L. Goldstone. 'Effects of interleaved and blocked study on delayed test of category learning generalization'. In: *Frontiers in Psychology* 5 (2014). ISSN: 1664-1078. DOI: 10.3389/fpsyg.2014.00936. URL: https://www.frontiersin.org/articles/10.3389/fpsyg.2014.00936.

[6] Mihaly Csikszentmihalyi. *Flow theory*. URL: https://www.sciencedirect.com/topics/psychology/flow-theory.

[7] Mouna Denden et al. 'Effects of gender and personality differences on students' perception of game design elements in educational gamification'. In: *International Journal of Human-Computer Studies* 154 (2021), p. 102674. ISSN: 1071-5819. DOI: https://doi.org/10.1016/j.ijhcs.2021.102674. URL: https://www.sciencedirect.com/science/article/pii/S1071581921000926.

[8] Ebbinghaus H. 'Memory: A contribution to experimental psychology.' In: ().

[9] Svenja Heitmann et al. 'Adaptive Practice Quizzing in a University Lecture: A Pre-Registered Field Experiment'. In: *Journal of Applied Research in Memory and Cognition* 10.4 (2021), pp. 603–620. ISSN: 2211-3681. DOI: https://doi.org/10.1016/j.jarmac.2021.07.008. URL: https://www.sciencedirect.com/science/article/pii/S2211368121000619.

[10] Joeri Dros Jaap M. J. Murre*. *Replication and Analysis of Ebbinghaus' Forgetting Curve*. URL: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0120644.

[11] Kahoot! *Kahoot!* URL: http://kahoot.com/.

[12] Sean H. K. Kang. *Spaced Repetition Promotes Efficient and Effective Learning: Policy Implications for Instruction*. URL: https://journals.sagepub.com/doi/full/10.1177/2372732215624708.

[13] Tuçe Köse, Eda ÿimen and Enisa Mede. 'Perceptions of EFL Learners about Using an Online Tool for Vocabulary Learning in EFL Classrooms: A Pilot Project in Turkey'. In: *Procedia - Social and Behavioral Sciences* 232 (2016). International Conference on Teaching and Learning English as an Additional Language, GlobELT 2016, 14-17 April 2016, Antalya, Turkey, pp. 362–372. ISSN: 1877-0428. DOI: https://doi.org/10.1016/j.sbspro.2016.10.051. URL: https://www.sciencedirect.com/science/article/pii/S1877042816312836.

[14] Teemu H. Laine and Renny S. N. Lindberg. 'Designing Engaging Games for Education: A Systematic Literature Review on Game Motivators and Design Principles'. In: *IEEE Transactions on Learning Technologies* 13.4 (2020), pp. 804–821. DOI: 10.1109/TLT.2020.3018503.

[15] Anton Lambers and Adrian J. Talia. 'Spaced Repetition Learning as a Tool for Orthopedic Surgical Education: A Prospective Cohort Study on a Training Examination'. In: *Journal of Surgical Education* 78.1 (2021), pp. 134–139. ISSN: 1931-7204. DOI: https://doi.org/10.1016/j.jsurg.2020.07.002. URL: https://www.sciencedirect.com/science/article/pii/S1931720420302361.

[16] M. A. McDaniel et al. *Test-enhanced learning in a middle school science classroom: The effects of quiz frequency and placement.* URL: https://psycnet.apa.org/record/2011-03466-001.

[17] Moodle. *Moodle.* URL: https://docs.moodle.org/401/en/About_Moodle.

[18] Moodle. *Quiz report statistics.* URL: https://docs.moodle.org/dev/Quiz_report_statistics.

[19] Tomoko Nemoto and David Beglar. 'Likert-scale questionnaires'. In: *JALT 2013 conference proceedings.* 2014, pp. 1–8.

[20] Sandra Y. Okita. 'Social Interactions and Learning'. In: *Encyclopedia of the Sciences of Learning.* Ed. by Norbert M. Seel. Boston, MA: Springer US, 2012, pp. 3104–3107. ISBN: 978-1-4419-1428-6. DOI: 10.1007/978-1-4419-1428-6_1770. URL: https://doi.org/10.1007/978-1-4419-1428-6_1770.

[21] Annemarie Sullivan Palincsar and Ann L. Brown. 'Reciprocal Teaching of Comprehension-Fostering and Comprehension-Monitoring Activities'. In: *Cognition and Instruction* 1.2 (1984), pp. 117–175. ISSN: 07370008, 1532690X. URL: http://www.jstor.org/stable/3233567 (visited on 23rd May 2023).

[22] Jay Parkes and Dawn Zimmaro. *Learning and Assessing with Multiple-Choice Questions in College Classrooms.* URL: https://www.taylorfrancis.com/books/mono/10.4324/9781315727769/learning-assessing-multiple-choice-questions-college-classrooms-jay-parkes-dawn-zimmaro.

[23] Quizlet. *About Quizlet.* URL: https://quizlet.com/mission.

[24] Stuart Sutherland Sandie Gay Michelle Bishop. *Chapter 8 - Teaching Genetics and Genomics for Social and Lay Professionals.* URL: https://www.sciencedirect.com/science/article/pii/B9780124201958000082#bib9.

[25] Gry Seland. 'Brukbarhetstesting av papirprototyper, TDT4180 våren 2006'. In: (2006).

[26] University of South Australia. *Evaluating Quiz Questions.* URL: https://i.unisa.edu.au/contentassets/a10bb28d309b43cfbcdec5d3c7d9d3c2/evaluating-quiz-questions.pdf.

[27] Scott W. Vann Andrew A. Tawfik. *Flow Theory and Learning Experience Design in Gamified Learning Environments.* URL: https://edtechbooks.org/ux/flow_theory_and_lxd.

[28] Ayanna K. Thomas et al. 'Should You Use Frequent Quizzing in Your College Course? Giving up 20 Minutes of Lecture Time May Pay Off'. In: *Journal of Applied Research in Memory and Cognition* 9.1 (2020), pp. 83–95. ISSN: 2211-3681. DOI: https://doi.org/10.1016/j.jarmac.2019.12.005. URL: https://www.sciencedirect.com/science/article/pii/S221136811930186X.

[29] TV2. *Dag Otto Lauritzen og Kristian Ødegård gjenforenes i sprøtt spørreprogram.* URL: https://www.tv2.no/a/6669303/.

[30] Kurt VanLehn, Randolph M. Jones and Michelene T. H. Chi. 'A Model of the Self-Explanation Effect'. In: *The Journal of the Learning Sciences* 2.1 (1992), pp. 1–59. ISSN: 10508406, 15327809. URL: http://www.jstor.org/stable/1466684 (visited on 10th Nov. 2022).

[31] Christopher N. Wahlheim, Geoffrey B. Maddox and Larry L. Jacoby. *The Role of Reminding in the Effects of Spaced Repetitions on CuedRecall: Sufficient but Not Necessary.* URL: https://psycnet.apa.org/fulltext/2013-28766-001.pdf.

[32] Alf Inge Wang and Rabail Tahir. 'The effect of using Kahoot! for learning – A literature review'. In: *Computers Education* 149 (2020), p. 103818. ISSN: 0360-1315. DOI: https://doi.org/10.1016/j.compedu.2020.103818. URL: https://www.sciencedirect.com/science/article/pii/S0360131520300208.

[33] DC: The National Academies Press Washington. '6 Designing an Assessment System'. In: *National Research Council* (2014). DOI: https://doi.org/10.17226/18409. URL: https://nap.nationalacademies.org/read/18409/chapter/8.

[34] David Sharek Eric Wiebe. *Using Flow Theory to Design Video Games as Experimental Stimuli.* URL: https://journals.sagepub.com/doi/abs/10.1177/1071181311551316.

[35] Piotr Wozniak. *Exponential nature of forgetting.* URL: https://supermemo.guru/wiki/Exponential_nature_of_forgetting.

Appendix

A Mind map

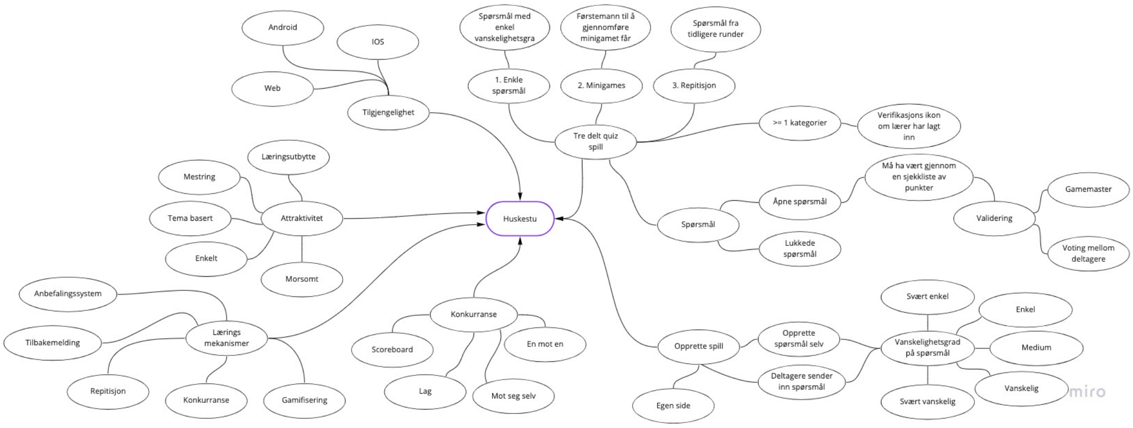


Figure 35: Mind map showing our initial thoughts

B User test post-interview

Spørsmål:

Game & Gameflow
Introduksjon:
Fikk du en forståelse av hva som skulle skje under introduksjonen?

-

Fikk du tid til å reflektere underveis i spillet?

-

Vanskelighetsgrad:
Hvordan var vanskelighetsgraden?

-

Følte du at spørsmålenes vannskelighetsgrad tilpasset seg deg?

-

Quizkombinasjon:
Hvordan var det å kombinere flere quizzer?

-

Spørring:
Var det lett å forstå hva som måtte gjøres underveis i spillet?

-

Var det lett å forstå hvem sin tur det var?

-

Hvordan var det å svare på åpne spørsmål?

-

Avstemning:
Hvordan var den demoktratiske avstemning for ritkig/feil svar?

-

Følte du deg sikker på når du stemte om et svar var riktig/galt?

-

Fikk du med deg hva andre svarte? (Fulgte du med når andre fikk spørsmål)

-

Fikk du diskutert spørsmål og svar med andre? Hvordan var det?

-

Spillfaser:
Hva syntes du om minigames? (Distraherende, motiverende, artig?)
-

Hva synes du om repetisjonen?
-

Sosial Dynamikk:
Hva synes du om lagbasert inndeling?
-

Hvordan påvirket andre spillere deg?
-

Feedback:
Forstod du konseptet med "forgetting curve"?
-

Var anbefalingene på hjemmesiden interessante for deg?
-

Er daily repetition noe du ville brukt?
-

Oppretting av Quiz:
Hvordan var det å velge kategori?
-

Hva synes du om tilbakemeldingene som errormeldinger, var de tydelige?
-

Svarte hjelpemodulene på det du lurte på?
-

Hvordan var det å sette vannskelighetsgrad og tid på svaret?
-

Brukergrensesnitt:
Hva synes du om designet på quizapplikasjonen?

-

Synes du designet til applikasjonen gjør den lettere eller vanskeligere å bruke? Hvorfor/hvorfor ikke?

-

Hvordan påvirket designet spillopplevelsen din? Ble det vanskeligere å følge med? Skapte det et mer lekent miljø?

-

Hvordan var det å navigere seg rundt på applikasjonen?

-

Har du noen forslag til hvordan applikasjonens design kan bli forbedret?

-

NTNU

Norwegian University of
Science and Technology