A Game Called Rufus
Designing A Syntax Driven Puzzle Game

Creating Level Design Strategies and
Comparing Their Impact on User Engagement

Bachelor's thesis in Computer science and engineering

Joel Jönsson
Johan Gottlander
Johanna Grahn
Mattias Retteli

A Game Called Rufus
Designing A Syntax Driven Puzzle Game

Creating Level Design Strategies and Comparing Their Impact on User Engagement

Joel Jönsson

Johan Gottlander

Johanna Grahn

Mattias Retteli

UNIVERSITY OF
GOTHENBURG

CHALMERS
UNIVERSITY OF TECHNOLOGY

A Game Called Rufus
Designing A Syntax Driven Puzzle Game Creating Level Design Strategies and
Comparing Their Impact on User Engagement

Mattias Retteli, Joel Jönsson, Johan Gottlander, Johanna Grahn

Cover: An image showcasing the game's main character in an environment drawn and inspired by the artist worked with.

A Game Called Rufus
Designing A Syntax Driven Puzzle Game
Creating Level Design Strategies and
Comparing Their Impact on User Engagement

Joel Jönsson
Johan Gottlander
Johanna Grahn
Mattias Retteli

Department of Computer Science and Engineering
Chalmers University of Technology
University of Gothenburg

Abstract

One way to evaluate whether a game is successful or not is to measure the engagement invoked in the player. This report presents an iterative design process focused on designing and implementing a syntax driven puzzle game. The purpose of this project was to answer these two questions: "What is the impact of different design strategies on the user's engagement while playing?" and "What should be considered when applying different level design strategies?". This was examined by developing a game, *A Game Called Rufus*, consisting of seven levels. These were designed using three different level design strategies created during the development process. The strategies were tested and evaluated using questionnaires and thematic analysis. The results showed some differences between the three strategies, but overall the experienced engagement were rather similar. It could not be determined with certainty how much the strategies impacted the result, as there were many other factors involved. However, the differences were further investigated and discussed. The strategies were also evaluated from a designer's perspective in order to present aspects that future designers have to take into consideration if working with these strategies.

Sammandrag

Ett sätt att utvärdera om ett spel är framgångsrikt eller inte är genom att mäta spelarens engagemang. I denna rapport presenteras en iterativ designprocess med fokus på att designa och implementera ett syntaxdrivet pusselspel. Syftet med detta projekt var att svara på frågorna: "Vilken påverkan har olika designstrategier på användarens engagemang när de spelar?" och "Vad bör finnas i åtanke när olika strategier för leveldesign appliceras?". Detta undersöktes genom att utveckla ett spel, *A Game Called Rufus*, bestående av sju nivåer. Dessa designades med hjälp av tre olika strategier för leveldesign som skapades under utvecklingsprocessen. Strategierna testades och utvärderades med hjälp av frågeformulär och tematisk analys. Resultaten visade vissa skillnader mellan de tre strategierna, men överlag var det inte så stora skillnader mellan det upplveda engagemanget. Det kunde inte fastställas med säkerhet hur mycket strategierna påverkade resultatet, eftersom många andra faktorer var involverade. Skillnaderna undersöktes och diskuterades dock vidare. Strategierna utvärderades också från en designers perspektiv med avseendet att presentera aspekter som framtida designers behöver ta hänsyn till vid arbete med dessa strategier.

Nyckelord: interaktionsdesign, pusselspel, Unity, leveldesign, spelutveckling, design-tänkande, handmålade illustrationer, barns ritningar

Acknowledgements

We want to thank our supervisor Mafalda Samuelsson-Gamboa, for all her help and support throughout this project. We also want to express gratitude to Michael Heron, for all his help regarding development in Unity. We would also like to thank everyone who has taken their time to contribute to this project by testing the game and giving us valuable feedback. Lastly, we would like to especially thank Rufus Forsman for letting us use his wonderful drawings in our game.

Joel Jönsson
Johan Gottlander
Johanna Grahn
Mattias Retteli
Gothenburg, June 2021

Contents

List of Figures

List of Tables

1

Introduction

Game design becomes increasingly popular, and with the help of free tools such as Unity [1] or Unreal Engine [2], almost anyone can create a game. However, in order to develop a successful game, it is worth researching what makes a game engaging. An essential part of achieving an engaging game is in its gameplay design, which is the process of conceiving, designing and implementing a game aiming at supporting specific feelings and experiences for the players. In addition to designing gameplay that reflects these feelings, it is important to design levels that let players experience the gameplay as intended. This is referred to as *level design* which is the part of a game development process that this report will cover.

In this report, the design process of the syntax-driven puzzle game "A Game Called Rufus" (AGCR) is described.

An example of an existing game with great level design is *Baba Is You* (BIY), which is a puzzle game released in 2019 [3]. The core mechanic of *BIY* is to move words in order to change the rules of the game. For example, removing the word "wall" from " wall is stop", manipulates the game rule making it possible for the character to move through walls, since walls no longer block the character. *BIY* is in many aspects similar to this project. It is also a syntax-driven puzzle game, but most importantly it is fun to play, keeping player engaged for hours. Because of how the mechanics work in this game as well as how it manages to engage the player, it became a key inspiration for this project.

Another influential game is *Gorogoa*, a puzzle game created by Jason Roberts [4]. *Gorogoa* integrates the player in a narrative told through hand-drawn illustrations. The gameplay is designed with four panels containing clickable illustrations that can be rearranged and connected in creative ways to further the game's narrative. Contrary to *BIY*, *Gorogoa* is based on hand-drawn illustrations and uses clickability (see section 2.1) as a core mechanic to advance in the game. Since the game developed in this project also is based on hand-drawn illustrations and uses a drag and drop mechanic, *Gorogoa* was a good source of inspiration.

1.1 Game Idea

A Game Called Rufus, the game being developed in this project, is a syntax-driven puzzle game based on preexisting drawings. These are drawn by an eight year old child called Rufus Forsman, hence the name of the game. Two examples of his drawings can be seen in Figure 1.1. The meaning of the game being syntax-driven is that the player is able to manipulate the rules of the game by moving words and creating different sentences on the display. An example of such a sentence is "Rufus can walk", implying that the player can move the character until the word "walk" is replaced by another verb; giving the character another ability.

Figure 1.1: Two of the drawings drawn by Rufus.

Usually, the final assets of a game are created at the end of the development process. However, since *AGCR* is based on already existing drawings, the environmental art assets are already finalized. Therefore, the game cannot be created using a traditional development process, requiring new level design strategies to be created.

1.2 Purpose

In order to research how player engagement is impacted by level design and how future developers can utilize this research, the purpose of the project is to answer the questions:

> 1. What is the impact of different level design strategies on the user's engagement while playing?
>
> 2. What should be considered when applying different level design strategies?

To answer these questions and fulfill the purpose, the process is divided into smaller goals to make it more comprehensible.

The first goal is to *design and develop* a game.

The second goal is to *create and evaluate* three level design strategies, and assess their impact on user engagement. The level design strategies are: "Level design through cooperation with an artist", "Level design based on existing environmental assets", and "Creating levels through puzzles". This will help answer the first question.

Finally, the third goal is to *compile the result of the evaluation.* This can be used as a resource to help future developers who wish to create a game based on preexisting drawings, and answers the second question.

2

Theory

This chapter presents theories and concepts used in this project. It introduces game design patterns, aspects of level design, and a model of player engagement and how to measure it.

2.1 Game Design Patterns

Game design patterns are descriptions of reoccurring parts of the design that concerns gameplay, and most of them are described by Björk and Holopainen [5]. This section introduces game design patterns relevant for this project with a brief explanation of what they mean, as well as an example of how they can be implemented in a game.

Clickability entails a visual element in game or in a menu/inventory that is clickable. Examples can be responsive buttons in a menu or items in an inventory system [6].

Destructible objects are object that can be destroyed in the game [7]. An example of a destructible object could be a box that has to be destroyed in order to proceed.

Enemies are game elements actively hindering the player from completing game goals. Examples of enemies are opponents in an first person shooter game [5, p. 70].

Goal hierarchies are structures that affect the possibility of a player to complete a certain goal, depending on if the player has completed other goals or not. A good example can be found in *The Legend of Zelda* [8], where the player gets new quests to complete throughout the game [5, p. 321].

Goal indicators are information that is presented to the player about their current goals in the game. Examples of the goal indicator pattern can be found in *The Legend of Zelda* [8], where the game instructs the player what to do next by tracking the current goal [5, p. 142].

Inventory is a space where the player can store game elements found in the game. An inventory system can be as simple as a grid in which the player can put items [9].

Levels are parts of the game in which all player action takes place until a certain goal has been reached, or an end condition has been fulfilled [5, p. 60]. A level can clearly be displayed in a menu with its own name, as in *Baba Is You* [3].

Scenes are fragments of a level. The gameplay can be divided into several different sequences differentiated by where and when they take place in the game world. Such individual sequences are Scenes [10].

Non-playable character (NPC) is a character in a game world that is not controlled by the player. An NPC is occurring within the context of the story whose purpose is to be interacted with by the player [11]. An example of an NPC is a character the player can talk to in *World of Warcraft*[12] to receive a quest.

Obstacles are game elements that prevents the player from taking the shortest route between two places in a game world [5, p. 75]. This could be a bottomless pit in a platformer such as *Super Mario Bros* [13].

Predefined goals are goals of the game that the designers have predefined before gameplay begins, meaning the winning condition [5, p. 310]. Goals could be save the princess or to kill a boss. A real world example can be found in chess where the predefined goal is to checkmate the opponent's king.

Puzzle solving are activities in the game that primarily can be solved through reasoning and problem solving [5, p. 368]. Examples of puzzles are memory games and jigsaw puzzles.

Single player games only involve one player in each game instance [5, p. 344]. An example is *Baba Is You*, where the player is the only person playing the game.

Storytelling means that there is one or multiple stories told within the game [5, p. 213]. A game that does this is *The Legend of Zelda* [8] series, where the player embarks on an adventure as Link.

Winning by ending gameplay implies that the player has won the game simply by reaching a certain game state which completes the gameplay [14]. A real world example can be found in chess where the game is over when one player checkmate the opponents king.

2.2   Level Design

Within gameplay design, level design refers to the process of creating a game environment the player can interact with. There are many aspects to consider when working with level design, such as the level design process and principles for good level design.

2.2.1 Level Design Process

Kramarzewski and Nucci present a general design process for level design [15, pp. 200-211]. The process begins with an idea which could be introducing a new mechanic or combining multiple existing ones into one level. Before implementing the idea into the game, a sketch is created to plan out the level. When the sketch is done and the team is satisfied with the vision of the level, implementation of the level can begin. Before any art assets are created, the level is implemented using greyboxing (see subsection 3.1.5). When the level design has been tested and the team is satisfied with it, art can be created and added to the level. The final step in the process is polishing, and consists of small changes to the level as well as bug fixes.

2.2.2 Principles for Good Level Design

Dan Taylor, a senior level designer working at Square Enix Montreal, held a session at the conference *Game Developers Conference* [16] in 2013. At this conference, Taylor discussed principles for good level design [17], the most relevant are described below.

Good level design is fun to navigate. Taylor claims that moving through a level can be broken down into three vectors; *observational*, taking in the level observing key points and obstacles; *strategic*, processing the information and making up a plan, and *navigational*, the act of traversing the level. Furthermore, the player should always know where to go which can be highlighted through either animations, colors, lighting, or geometry.

Good level design does not rely on words. A game's narrative can be divided into three categories according to Tyler; *explicit*, *implicit*, and *emergent*. Explicit narrative is any peace of information that is explicitly called out to the player, for example spoken or written dialogue, cut scenes, or text. Implicit is information the player figures out from the environment and can be done through mise-en-scéne [18]. Emergent is anything the player makes up in their mind.

Good level design tells what to do but never how to do it. Telling the player what to do without also telling how to do it is hard. Taylor suggests using a technique called *nebulous objectives* which entails giving the player short and concise information, only enough to make the objective clear. He also makes it clear that the designer should give the player a few alternative methods of completing the task and not punish the player for improvisation.

Good level design constantly teaches. Each level should either introduce, showcase, or subvert a key mechanic in the game to keep the player interested.

Good level design is efficient. When developing a game the resources are not unlimited, so it is up to the designer to make the most out of the given resources. A modular design helps with this, it works by stringing together a series of modular mechanic driven encounters that can be put together to make a level. These mod-

ules can then be reordered to create another level and this can be repeated giving the designer different levels from one design. This also reinforces the mechanic in the player by giving them time to practice it.

Good level design creates emotion. All art create emotions and so does video games as well. However, the tricky part is to create the intended emotions. A designer can play with the environment to create a desired feeling, and to know what parts of the environment to change the designer can work backwards; starting with a desired emotional response and then deciding the spacial parameters, narrative elements, and mechanics to achieve that emotion.

Good level design is driven by mechanics. A level should be a gameplay-delivery system to leverage the mechanics to create a great experience [16]. Quests, combat, story, environment should all showcase the mechanics.

2.3 Engagement

This section will briefly cover the model of engagement presented by O'Brien and Toms [19], as well as a description of the revised user engagement scale [20] which also is based on their work on measuring engagement [21].

2.3.1 Model of Engagement

In 2008, O'Brien and Toms published an article with the purpose to "critically deconstruct the term *engagement* as it applies to people's experiences with technology" [19, p. 938]. They propose a model with four stages of engagement. These are *point of engagement*, *period of engagement*, *disengagement*, and *re-engagement*. To engage a user, there must be a point of engagement. This could be an aesthetic appeal, interest or a goal. Period of engagement refers to the user's thoughts and feelings while engaged, as different factors as to why they stay engaged. A user can become disengaged if they for example feel like they have completed their task or are frustrated with the product. If there is nothing to re-engage the user again, they might stop using the product.

2.3.2 Revised User Engagement Scale

O'Brien and Toms [21] presented a survey with 31 items divided into six factors to measure user engagement. These six factors were *focused attention*, *perceived usability*, *aesthetics*, *endurability*, *novelty*, and *felt involvement*. A Likert scale (see Figure 3.2.5) was used to measure each item. The scale consisted of five points, with "strongly disagree" and "strongly agree" being at the opposites endpoints. There was an option to not answer a question if the user did not feel like they could answer it. Also, the order of the items in the survey were randomized.

Wiebe *et al.* [20] revised O'Brien and Toms' user engagement scale to better relate it to video games. They did this by merging the three factors *endurability*, *novelty* and *felt involvement* into one factor, and removing three items because they lacked correlation with other items. The new factor was called *satisfaction.*

Focused attention focuses on user awareness, both inside as well as outside of the user's experience in the product. Awareness focused on the experience in the product could be how involved and absorbed the user is. The lack of awareness on things around the user and how disconnected the user is from the world are also taken into account. For example, the user might become solely focused on the task in the product, and things in the real world might go unnoticed [20].

Perceived usability is described as "Overall, these items assessed users' perceived effort in using the Website, their ability to accomplish their shopping tasks, the navigation and organization of the Website, and the emotions evoked by using the Website" [21, p. 56]. Perceived usability contains items that consider the emotions connected to the difficulty of using the product. It also considers the source of these emotions. For example, if the task took too much effort or was too challenging for the user, it can incite frustrations or discourage them.

Aesthetics focuses on the visual aspect of the product. Items that asses the product's visual appeal and attractiveness are grouped in this factor. Both graphics and layout are considered [21]. This factor should not be confused with other uses of the word "aestethic", which considers more aspects than only visuals, such as for example seen in the MDA model (see subsection 3.1.4).

Satisfaction is the result of merging the three factors *endurability*, *novelty*, and *felt involvement.* The factor is a measurement of how fun the experience was, if it incited curiosity, and if it was interesting. It also considers if the user would want to use the product again and recommend it to others [20].

3

Method

This chapter contains a description of the methods used in the development process. It starts with introducing the software used for this project, followed by an explanation of the design process, and a clarification of the chosen working methods. Lastly, the methods used for user testing and level design strategies are presented.

3.1 Tools

This section presents software used to execute this project. It also covers the MDA framework used for level design analysis, and greyboxing used for level prototyping.

3.1.1 Figma

Figma is an interface design application developed by Figma, Inc [22]. It is available to use online and is very useful when collaborating since multiple users can work on the same project with real-time updates. This tool facilitated the work of the project since all work was done from home because of the ongoing pandemic. Figma was used to create low- to mid-fidelity prototypes which resembled paper prototypes in a digital format.

3.1.2 Unity

Unity is a game engine developed by Unity Technologies [1]. A game engine is an environment for software development which allows the user to build video games. It contains many built-in tools to facilitate game developing and to make it available for users with various levels of programming knowledge.

3.1.3 Git

Git [23] is an open source distributed version control system designed to operate projects fast and efficiently. Using Git made it easier for multiple people to work on the same codebase without causing unnecessary errors.

3.1.4 MDA Framework

The MDA framework is a term used in game design and is a tool used to analyze games. The abbreviation MDA stands for "Mechanics-Dynamics-Aesthetics"; *mechanics* describes the rules of the game, *dynamics* describes how the game is played,

and *aesthetics* describes the feelings that should arise while playing the game [24].

3.1.5   Using Greyboxing in Level Design

Greyboxing is a method used to quickly create a prototype or a playable experience through use of primitive assets [25][26]. As the name suggests, the assets often consists of grey boxes, or shapes without any texture (see Figure 3.1). With this technique, a designer can quickly create a level and get a feel for the perspectives, environment, and scale. The technique allows testing in order to make sure that the assets work as intended without having to put in massive amounts of time to create each asset.
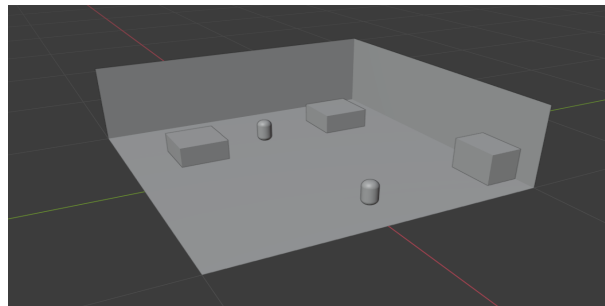
Figure 3.1: An example of greyboxing in a 3D game.

3.2   Design Thinking Process

Design thinking (DT) is a methodology used in design that provides a solution-based approach when it comes to solving difficult problems [27]. These problems can be described as wicked problems. A wicked problem is a problem with many interdependent factors, making them impossible to solve as there are no defined solutions. Many design problems are wicked problems, which means that clarifying the problem itself generally is as big a task as finding a solution [28]. Since there is no criteria for when a solution has been achieved, this also means that there is no stopping rule to a wicked problem. To tackle this, the DT process can be iterated indefinitely, but eventually has to be stopped. It might be stopped due to time limitations, funding, or simply because the current solution is good enough [29].

The DT process requires an understanding of the human needs involved by reframing the problem in a more human-centered way, by proposing many ideas for solutions, and by embracing a hands-on approach with prototypes and testing. A result close to a solution is then achieved by iterating these steps until a product that satisfies the task has been decided on [27].

According to the Hasso-Plattner Institute of Design at Stanford, the five stages of DT are: *Empathise*, *Define* (the problem), *Ideate*, *Prototype*, and *Test* [27]. The

five stages are rarely sequential, as they do not have to occur in a particular order. More often than not, it is beneficial to work through the stages in parallel or repeat them iteratively with new results or ideas. However, the stages will be described as a sequential order below.

3.2.1 Empathize

Empathize is first stage of the DT process. At this stage designers gain sympathetic understanding of the problem in question. This can be achieved by consulting experts, or through engaging and empathizing with the target group, which makes it easier to understand their experiences and needs. Empathy is essential in a human-centered design process and allows the designers to set aside their own assumptions to gain knowledge about the users and what they need. In conclusion, the aim of this stage is to gather information to use in the next stage and to develop a better understanding of the users, including their needs and problems [27].

3.2.2 Define

The second stage of the DT process is *define*. The information collected from the previous stage is now used for analyzing what the core problems are, as well as formulating them as problem statements [27]. A possible method for the analysis is "Thematic Analysis".

Thematic analysis is used for analyzing qualitative data, such as interviews. The data is analysed by identifying common patterns that are then represented by themes [30, p. 322]. Similar data, for example data that express the same emotion or thought, are grouped together. These groups of data can then be combined or inserted into a theme. Themes can be chosen from two different approaches, *inductive* or *deductive*. The inductive approach allows the themes to emerge from the data collected, while the deductive approach is more about fitting the data into existing themes [31].

3.2.3 Ideate

Ideate is the third stage, where the designers generate ideas based on the information gathered and defined in the two previous stages. At first, it is important to come up with as many ideas and solutions as possible [27], and two techniques that can help with this are *brainstorming* and *sketching*. Brainstorming is an efficient method for generating ideas, increase creativity, and finding solutions. The task is to generate many different ideas or solutions to a problem, without criticism or limitations [32, p. 2]. Sketching is a good technique to explore the design space in a quick and inexpensive way [33], by visualizing ideas and potential solutions early on.

3.2.4 Prototype

The fourth stage is the *prototype* phase. This is when a number of inexpensive, scaled-down versions of the product or particular features found within the product is made. The aim of prototyping is to pinpoint the best attainable solution for each problem detected during the former stages. The found solutions are implemented in the prototypes and will be individually inspected and processed until the features are either accepted or rejected on the basis of the users' experiences [27]. The extent of how accurately a prototype represents the appearance and interaction of the actual product is the main determining factor when it comes to prototype fidelity [34, p. 78].

Low-fidelity prototypes are useful for depicting concepts, design elements, and layout. They can consist of static windows and elements that can be quickly generated. In general, a low-fidelity prototype has limited or no functionality. The users can see what the product is supposed to do but the system response may need to be simulated by a human. The purpose of a low-fidelity prototype is to illustrate the general structure of an interface, such as placement of elements, as well as provide a feel of how the application flows. This is done in order to collect and analyze the product in the early stages of the design process [34, p. 79].

High-fidelity prototypes as opposed to low-fidelity prototypes, contain more functionality and are interactive. When using a high-fidelity prototype, the user has the possibility to operate with a system much closer to the final product, compared to a low-fidelity prototype. For instance, the user would be able to press buttons on the screen and expect corresponding actions to execute [34, p. 81]. A high-fidelity prototype is used in the later stages of the design process, but does not equal the finished product.

3.2.5 Test

Testing is the fifth and final stage of the DT Process. This is the stage where the prototypes are put to test by several test users from the target group, in order to receive feedback so the product can be further developed [27]. Evaluation centers around the usability, how easy it is to learn and use, but also the users' experience when interacting with the system. For example, how enjoyable, motivating, or satisfying the interaction is [30, p. 496]. In order to collect and analyze collected data and feedback, the following methods can be used:

Think-aloud technique is a protocol used to understand the users' actions and examine their problem-solving strategies [30, p. 296]. It is an efficient technique to get insight of what is going on inside the user's head by asking the users to narrate how they are thinking as they execute actions with the prototype. The disadvantage of this protocol is that the narration may impact the experience.

Controlled setting involving users is a type of evaluation where the user is prompted to resolve specific tasks instead of letting the user explore freely, usually in a controlled setting such as a lab. One method which is often used in this setting

is *usability testing*. This is a good method for when "the goal is to test whether the product being developed is usable by the intended user in order to achieve the tasks for which it was designed and whether users are satisfied with their experience" [30, p. 524].

Questionnaires are a form of data gathering. They can be used to efficiently gather large amount of data. Questionnaires can be used together with other data gathering methods to get a better understanding of the situation [30, p. 278].

Semantic differential technique allows for degrees of opinion when answering questions in for example a questionnaire. This is done by having the respondent selecting a position on a bipolar adjectival scale [35]. The two ends of the scale host antonym adjectives (e.g. easy - hard), and the respondent selects the position which they feel corresponds to the experience (see Figure 3.2).

Figure 3.2: An example of a question using the semantic differential scale.

Likert scale allows for measurement of agreement in questionnaires [35]. The respondent is presented with a number of options that represents how much they agree or disagree with a statement. The opposite endpoints represent strong disagreement and strong agreement, with the midpoint being a natural stance on the statement.

Unstructured interview is another method used to gather data. Unlike questionnaires or structured interviews, the questions asked and discussed can vary a lot between the interviews. Interviews are more time consuming than questionnaires but can allow for greater understanding [30, p. 269].

3.3 User Tests

This section covers the general approach for the user tests conducted in this project. The three iterations were the *Pilot evaluation*, *Unity test* and *Final evaluation*. The process of the tests could be divided in four different steps: (1) planning, (2) implementation, (3) testing and (4) data analysis. This will be the structure used to describe the user tests in chapter 4.

All tests were set up in a controlled setting. Because of the ongoing pandemic, most of the tests were conducted digitally through video calls. The tests were conducted by one evaluator since it was possible for one person to instruct and document the tests. After the test, the participant filled in a questionnaire. Also, in order to make

sure that the necessary data could be collected and analyzed with the permission from the participant, a consent form was created that had to be signed by the participants before the test.

3.4 Level Design Strategies

One goal of this thesis was to create and compare different strategies that can be used when designing a puzzle level. The strategies were created based on a general design process with slight alterations in some areas. Since most of the art were already created from the beginning, one of the last steps in the general design process was done before any actual work on the level had begun.

Furthermore, when a level is to be constructed there needs to be an idea of the theme. This theme should be kept in mind and central when designing the level. Examples of themes could be a central asset such as a box or a mechanic such as "push". The created strategies have some core differences, however the general design process for the strategies are similar (idea, sketch, prototype, implementation, additional art, polish), but it is the idea stage that really separates them from each other.

A helpful technique when designing levels is reverse engineering. Traditionally reverse engineering, sometimes called back engineering, is described as a process in which software, machines, architectural structures and other products are deconstructed to bring out design information from them. To us, a reverse engineering design process would mean working backwards from the solution to the start of the puzzle. This is a useful technique to create puzzles or prototypes of puzzles without having a clear idea of how it should look in the end. This is done through defining how the end state of the level would look. For example where the goal is and where the player is, maybe even the immediate steps before the player reaches the goal as well. When the goal is clearly defined the developer imagines how the player gets there, for example a bridge or jumping on top of a box. When the developer settles on a route to the goal it is time to introduce puzzles to hinder the player form just walking up to the goal. After the first puzzle is introduced another one is created however this time making sure that this puzzle does not interfere with the previous. This is then repeated until the developer is satisfied with the length of the level and the difficulty.

4

Development

This chapter describes the development process of the game, which is represented as a timeline in Figure 4.1. Each stage of the timeline will be written about in its own section. The process starts with an outline of the initial research stage which focuses on the progression of game ideas. The following sections describes the process of creating prototypes and performing the two user tests conducted in order to gather data for the progression of the work toward the final prototype. After each iteration there is also a section describing each of the three level design strategies which were created to solve discovered problems. Finally, there is a section describing the final prototype and how the final evaluation was planned and analyzed.
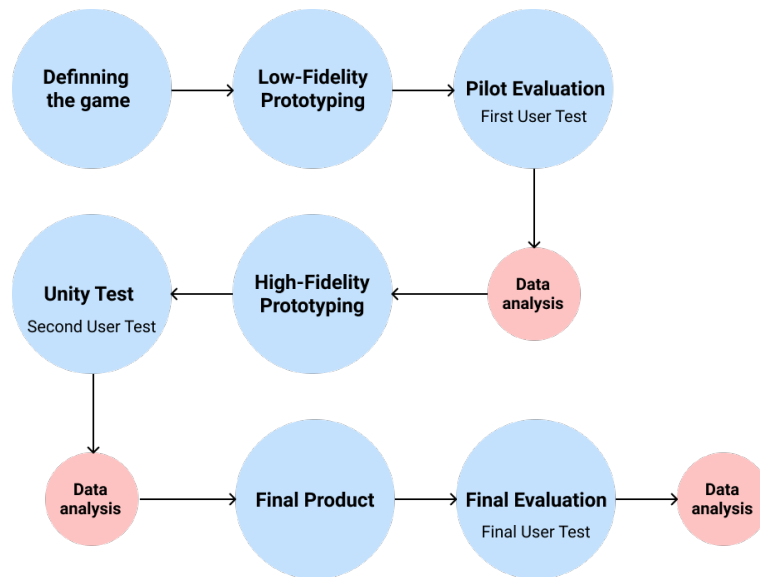
Figure 4.1: The development process represented as a timeline, from top left to bottom right.

4.1 Defining the Game

One of the first steps to define the rules and mechanics of the game was to research game patterns, the MDA framework, and game development through an interaction design process. Also, research into the development of puzzle games as well as how other puzzle games use different mechanics, dynamics, and aesthetics. Two relevant and inspiring games are *Baba Is You* and *Gorogoa*. Playing other puzzle games

helped inspire new puzzle ideas. Parallel to this, different potential ideas for the game were sketched.

4.1.1 Sketching

To come up with potential game mechanics, all members made several sketches (see Figure 4.2) based on their ideas and visions of the game, and on the analysis of similar games. The sketches were discussed and then refined to get a grasp of the game's goals considering mechanics and aesthetics. When this was done and a unanimous vision had been established, a moodboard (see Figure 4.3) as well as a MDA map were created.

Figure 4.2: Three examples of different sketches in early development.
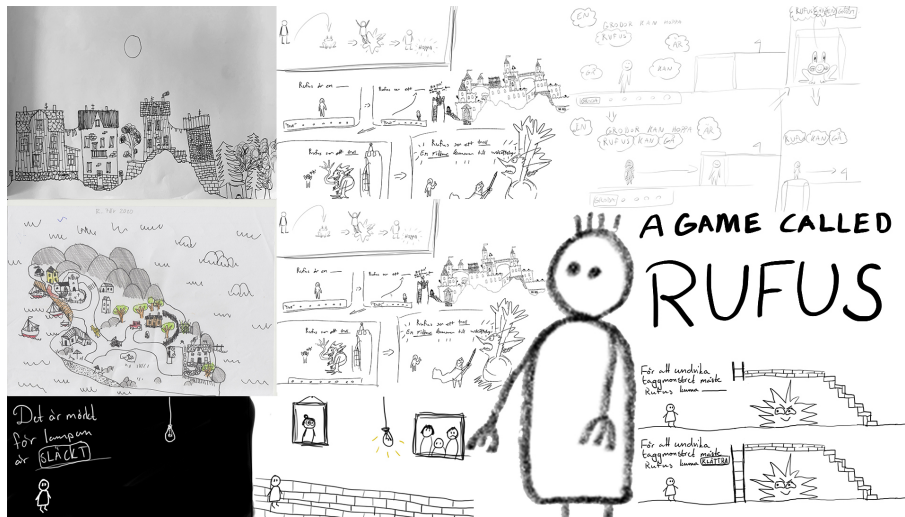
Figure 4.3: Sketches were inserted into a moodboard that was a first attempt to visualize how the gameplay would look like when it comes to the aesthetics.

4.1.2 MDA

The MDA framework was used to define the game and to make sure that the project would move forward in a direction aligned with the intended vision and feel of the game. The three components of the MDA can been seen in Figure 4.4, but are described in text below.
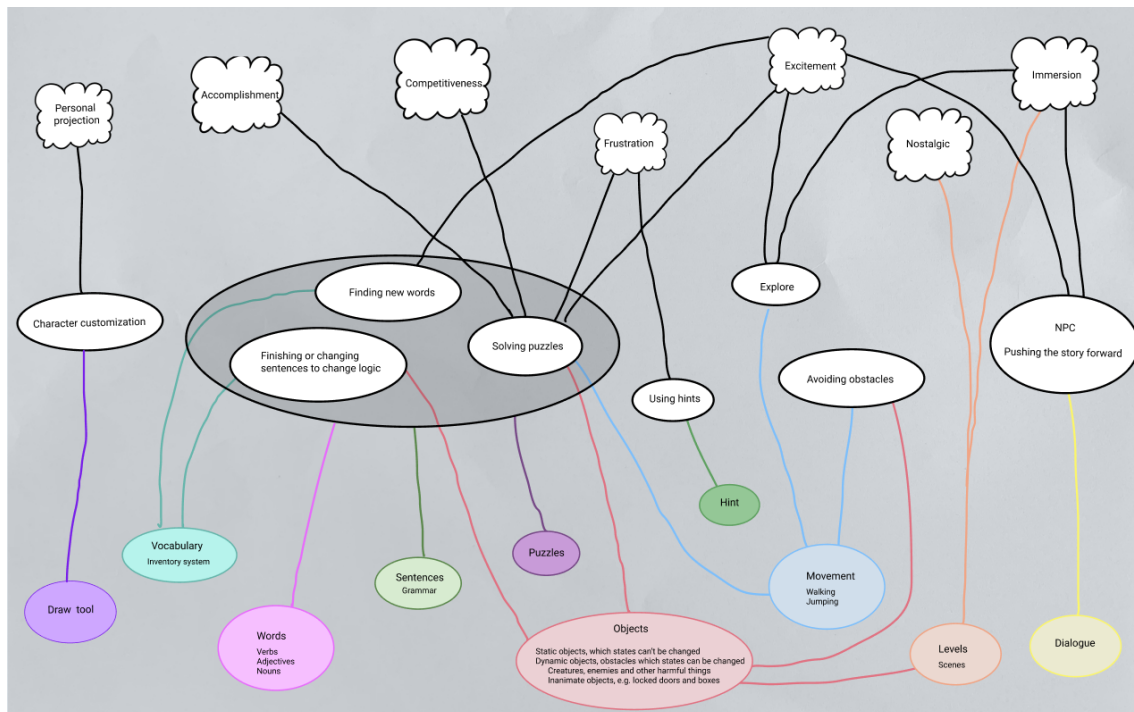
Figure 4.4: The MDA in a mind map style.
Mechanics at the lower level, Dynamics in the middle and Aesthetics at the top level.

Mechanics

- Movement such as "walking", "jumping", and "pushing".
- Sentences and words, which are fundamental for the player to be able to manipulate the rules of the game.
- Vocabulary, an *inventory* (see section 2.1) to store the words collected in the game.
- *Levels* (see section 2.1) that are built up by a number of puzzles or problems to solve.
- *Obstacles* (see section 2.1)
    - Static objects, whose state cannot be affected by creating new sentences.
    - Dynamic objects, obstacles that can be changed with new sentences.
    - Creatures that behave as enemies and other harmful things.
    - Inanimate objects, such as locked doors and boxes that the player cannot pass through.
- Dialogue that pushes the story forwards and helps the player solve a puzzle.
- *Puzzles* (see section 2.1) that stops the player from progressing in the game and demands the player to use problem solving skills to move forward.

Dynamics

- Finding ways around obstacles to progress in the game.
- Finding new words to use for manipulating the rules of the game.
- Customizing the character to let the player add a personal touch to the gameplay through personal projection.
- Interacting with non-playable characters to get hints for the puzzles and push the story forward.
- Move words from sentences or inventory to complete or change sentences.
- Finishing or changing sentences to change logic which will allow players to make different shifts in the game.
- Puzzle and problem solving which the player will have to use to complete the game.

Aesthetics

- Sensation
    - Should feel like playing through a child's drawings in order to bring a nostalgic and playful feeling to the player.
    - Rewarding, both in learning purposes and enjoyment.
- Narrative
    - Immersion and investment in the story and the characters.
- Challenge
    - Excitement when the player learns something new or is faced with a new obstacle to overcome.
    - Frustration when a problem seems unsolvable.
    - Sense of accomplishment after finally solving a tricky puzzle.

    – Competitiveness. If the game is played with a friend or in school, a sense of competitiveness can be emerged and push the player to try and solve the problem quicker than the other person.
- Expression
    – Personal projection, accomplished when the player can customize name and face to look like themselves. Makes the player feel a closer connection to the avatar they are playing as.

4.2 Creating Strategy "Level Design Based on Existing Environmental Assets"

Before the project started some assets had already been created. These were environments such as a city, a castle, and a boat. Each environment had interesting objects in it that could be used when creating obstacles in the level. From this, three steps were created to further refine the strategy, and to make it more accessible. These are further described later in subsection 5.3.1.

4.3 Low-Fidelity Prototyping

Firstly, several digital paper prototypes were created based on the refined sketches. This was done in order to encompass all aspects of the game, and outline a common vision for game. The prototypes featured a number of small-scale levels. These were low-fidelity prototypes, and contained little functionality, but were easy and quick to make. The benefit of creating such prototypes was to get a feel of how different game mechanics work in comparison to each other.

4.3.1 Prototyping in Figma

The tool "Figma" [22] was used to create several digital paper prototypes, with the goal to be merged into one final low-fidelity prototype. These prototypes were envisions of how the different game mechanics could be combined to create puzzles, and an example can be seen in Figure 4.5. At this point, the general layout of the game had been set which included an inventory of words at the bottom of the screen.
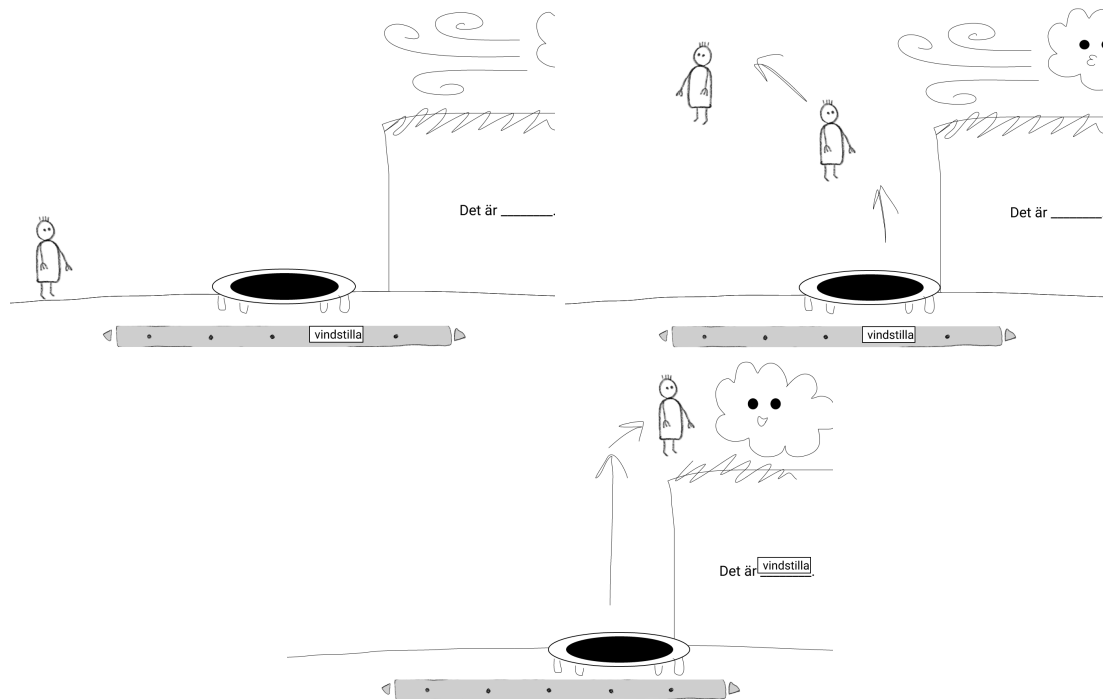
Figure 4.5: A digital paper prototype made in Figma, showcasing the main mechanic of the game.

When enough low-fidelity prototype levels had been made, they were summarized into one small prototype in Figma which will be shown in section 4.5. This was used in the first user test and together with the feedback gathered, it became the foundation of a first high-fidelity prototype further developed in Unity.

4.3.2 Problems Discovered During Prototyping

The first attempts on designing levels resulted in prototypes that focused on narrative and mechanics, as in character movement and changing sentences, lacking an actual puzzle to solve. The designs looked more like an interactive story rather than a puzzle game (see Figure 4.6). When creating new prototypes with this in mind the puzzles were improved. However, there were difficulties finding enough elements from the drawings that could be used. This became a problem since the aim was to keep the aesthetic of drawings made by a child.
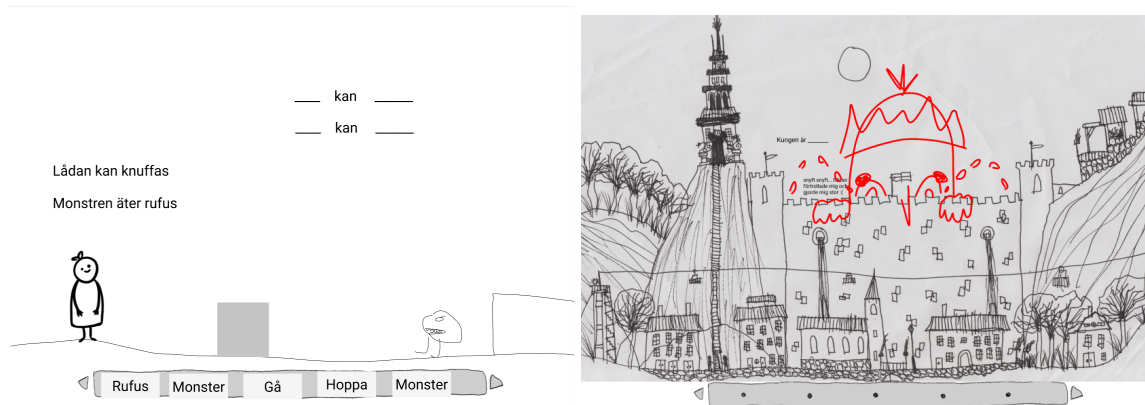
Figure 4.6: A puzzle based prototype (left) and a frame from a story based prototype (right).

4.4 Creating Strategy "Creating Levels Through Puzzles"

During the prototyping phase it became apparent that some drawings were better suited for the level design strategy "Level design based on existing environmental assets" than others. Furthermore, the created levels were too linear and did not give enough of a challenge. As an attempt to solve this, puzzles were created without considering the environment. However, they were too short and therefore intertwined with other puzzles in order to create longer levels.

4.5 Pilot Evaluation

The user tests were carried out in Figma using the low-fidelity prototype. The goal of the evaluation was to see how a user would interact with the game and if there were any disconnects between expected and actual experience.

Step 1: Planning

The aim with the pilot evaluation was to observe users and get a perception of how they proceed to solve a level. Information about the game's difficulty, time to complete levels, how difficult the mechanics were to grasp, and if the experience was enjoyable were also data that was collected. This evaluation focused on the mechanics of the game, rather than the design and environment. This was important because in order to continue implementing features in Unity, the core mechanics had to be figured out. The design can always change and is easier to adjust over time, but the core mechanics must be determined quite early.

In order to collect data about the participants' feelings and impressions about the prototype, a questionnaire was created using the *semantic differential technique*. As seen in Figure 3.2, one question was if the users perceived the mechanic involving

the drag-and-drop feature of the words, worked good or bad.

Step 2: Implementation

The prototype used for this test was a digital paper prototype created in Figma. Puzzles that introduced the game mechanics and represented them well were chosen to be part of the first evaluation, and the four levels of the prototype are shown in Figure 4.7).
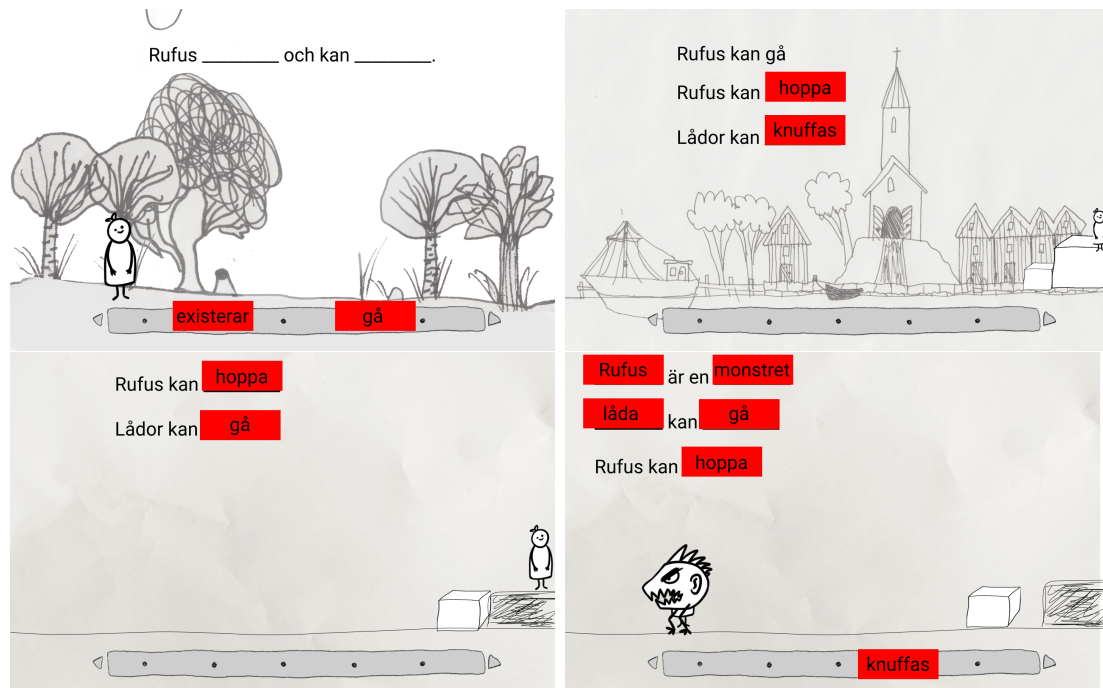
Figure 4.7: The low-fidelity prototype, containing four levels tested in the Pilot Evaluation.

Step 3: Testing

Participants were asked to interact with the prototype in Figma, and since the prototype could not respond to user input, the evaluator "acted as the computer", as in moving the character when the user wanted to do so. Observers documented comments and actions taken by the participants. After the participant had played through all four levels, they were asked to fill out a questionnaire about their experience which is presented in the next step. After the participant had filled in the questionnaire, they had the opportunity to talk freely about their thoughts on the prototype. The observer could also ask follow-up questions based on events from the game or questionnaire.

Step 4: Data Analysis

The data from the pilot evaluation was gathered and analyzed in different ways to help us understand the users' experience of the prototype. The first step was to

analyze the data from the questionnaire by calculating the average value of all the answers for each question, where each question could have an answer within the range 0-7. The results can be seen in Table 4.1

Statement		Mean
I think the game was:	Boring - Fun (0-7)	5.4
While playing I felt:	Bored - Engaged (0-7)	6.2
While solving puzzles I felt:	Frustrated - Relaxed (0-7)	4.6
While playing I felt:	Concentrated - Distracted (0-7)	1.6
At the end I felt that I had:	Failed - Performed (0-7)	6.2
In terms of time, the game was:	Too short - Too long (0-7)	2.8
I think the puzzles in the game were:	Simple - Tricky (0-7)	3.8
I felt that the mechanic (drag words) worked:	Bad - Good (0-7)	6.2

Table 4.1: Questionnaire results from the Pilot Evaluation, showing the mean for each statement.

The next step was to analyze the notes written down and the observations made during the user tests. This was done by identifying different themes using an *inductive* thematic analysis [30, p. 322]. Inductive reasoning aims at developing a theory.

The first step of the inductive analysis was to collect feedback from the users (see Figure 4.9). To separate each user's feedback a color scheme was used, where each color represents a specific participant. The next step was to group data that were connected into collections. When this was done, the last step was to come up with a theme that described each collection of feedback. This was done in two levels, where one overlaying theme described a collection of sub themes. The three main themes were "Feelings", "Approach" and "Uncertainties". These were then divided into smaller themes, where "Feelings" for example was divided into "Accomplishment", "Pleasure", and "Relaxed - Frustrated".
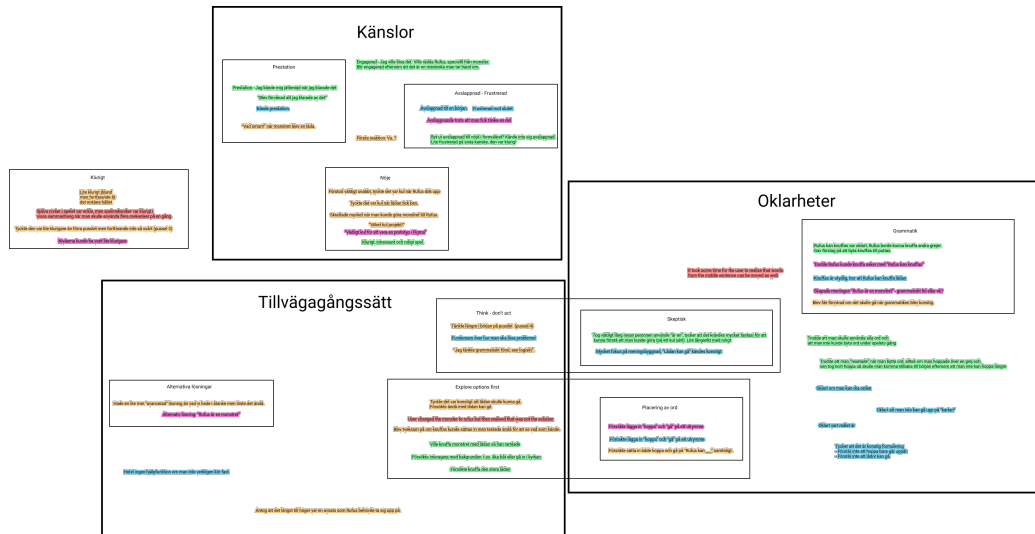
Figure 4.8: Inductive analysis of the pilot evaluation

4.5.1 Conclusions

Feelings. The data from the feedback shows that the majority of the users felt that they had accomplished something while also having a good time interacting with the prototype.

> *"I felt very pleased when I made it."*

> *"Very fun considering it is just a prototype in Figma."*

Approach. The analysis of the different approaches showed that users had various methods of completing a level. Some tried to explore the different options first before acting, while others acted before thinking. This showed that it is important to let the player play the way they want to, and not be forced to play a certain way.

> *"I thought grammatically first, then logically."*

4.5.2 Problems Discovered

There were a lot of uncertainties which mostly had to do about grammar and the placement of words. The biggest issue was the users perception of the Swedish word "knuffas" since the way this word was implemented in the game did not correspond to how the users would use the word themselves. The users wanted to build the sentence "Rufus kan knuffas" in order to give Rufus the ability to push other objects, but this had no effect when they tried this. This is because the implementation had the reverse effect where the sentence "Rufus kan knuffas" implies that Rufus can be pushed by other objects. So in order to be able to push a box you would have to make the sentence "Lådan kan knuffas", making it possible for the box to be pushed by Rufus.

"I thought Rufus could push other things with "Rufus kan knuffas"."

Apart from changing "knuffas" to "puttas" based on user suggestion, no other major changes were made. The team was happy with the results and it seemed like the development was going in the right direction.

4.6 High-Fidelity Prototyping

The creation of a Unity prototype started off with constructing one scene with some simple objects that were used to test the functionality of the code. Once that had been finished, scripts for some simple mechanics of the gameplay were written. Examples of such scripts are the ones for player movement and dragging the words around the screen with the cursor. As a next step, a series of scripts defining the core idea of the game were created. These scripts handled the sentence building aspect of the game. When a word was added to a sentence, the scripts applied the functionality of the word to the object pointed to in the sentence. This approach created an unintentional bug where words would inherit the same functionality they were representing. For example, when applying the word "walk" to a sentence, not only the object in the sentence got the ability to walk, but also the word itself. This was solved by separating the word in the game world and the functionality of it. When the core functionality of the game had been created, levels and art assets could be added. The assets were either parts taken from the finished pictures that were gifted for the project from the start, or drawn by ourselves when new assets were needed. When the assets had been made and implemented in Unity, animations for walking, jumping and other moving parts were drawn and implemented as well.

4.7 Unity Test

When the pilot evaluation had been performed, the next step was to implement the levels into Unity so a user test with a real build of the game could be conducted. The aim of this test was pretty much the same as the aim of the pilot evaluation, but now the test user would play an early build of the actual game rather than interacting with a digital paper prototype. This was done in order to get better data on the mechanics and the feel of the game as well as catching new errors and uncertainties. The levels were the same as in the pilot evaluation, except that level three and four were merged into a single level and one extra level was created.

4.7.1 Test Structure

The second evaluation followed the same structure and execution as the pilot evaluation, with the only difference being the prototype. It was set up in a controlled setting (see subsection 3.2.5) and the questionnaire was the same.

4.7.2 Data Analysis

Below is the result from the questionnaire (see Table 4.2) with the corresponding average answer, and the inductive analysis technique was used to establish themes from the user data (see Figure 4.9).

Statement		Mean
I think the game was:	Boring - Fun (0-7)	5.4
While playing I felt:	Bored - Engaged (0-7)	5.6
While solving puzzles I felt:	Frustrated - Relaxed (0-7)	4.2
While playing I felt:	Concentrated - Distracted (0-7)	1.2
At the end I felt that I had:	Failed - Performed (0-7)	4.8
In terms of time, the game was:	Too short - Too long (0-7)	2.4
I think the puzzles in the game were:	Simple - Tricky (0-7)	4.6
I felt that the mechanic (drag words) worked:	Bad - Good (0-7)	6.4

Table 4.2: Questionnaire results from the Unity Test, showing the mean for each statement.

Figure 4.9: Inductive analysis of the data gathered from the Unity test.

4.7.3 Problems Discovered From User Feedback

A large portion of the data was collected in the "confusion"-theme, where all uncertainties and confusions were represented. Unfortunately, there was still a lot of confusion around the word that makes an object "pushable", even after it was changed from "knuffas" to "puttas". Another problem discovered was that there was no message showing the user that they were stuck, and they had to guess if a restart was required or if they should keep trying. Furthermore, the lack of a proper tutorial teaching the player how to control the character caused some participants not being able to progress through the level by themselves. There where also a lot of uncertainties around graphical elements. Most of them were about old assets that had not been replaced yet or art that was still a work in progress. Some participants commented on levels not being clearly separated as it was not obvious when a level started and ended.

4.8 Creating Strategy "Level Design Through Cooperation With an Artist"

At this point much of the already pre-made assets were used in other puzzles, there were also some problems with fitting puzzles to the unused assets. It was then tried to create the puzzles in a "void" as in the *creating levels through puzzles* strategy. The puzzle where then tested to make sure it was playable and interesting. After it was deemed so, Forsman was notified and asked to create an image based around the puzzles, and incorporate them in to the environment.

4.9 Final Prototype

The final game uses the Unity prototype as the core but with modifications according to the received feedback. Based on the data analysis, several changes were made. To improve clarity behind the meaning of "be pushed", the word "puttas" was changed to "flyttas". "Flyttas", meaning "be moved", has a similar meaning in this context. However, "kan flyttas" cannot be interpreted as "can move (other objects)", which makes it more clear than "kan puttas", that could be interpreted as "can be pushed" as well as the unintended "can push (other objects)". Using the "Good level design constantly teaches" principle (see section 2.2) a new level with the single purpose of introducing the word "flyttas" was created. Hopefully, these two changes combined will remove any confusion surrounding the word. This newly created level will only serve as a tutorial of sorts and will not be included in the design strategies analysis.

To help less experienced users, a help system was introduced, which in the game is referred to as "the language cop". His job was to clearly show the player when grammar errors occurred (see Figure 4.10). He also had a secondary purpose as to introduce the movement mechanics when appropriate. This was achieved by waiting a few seconds to see if the player would move before telling them how to do

it. Furthermore, a lot of places where the player could get stuck were removed or changed with the intent to not to force the player to restart the level.
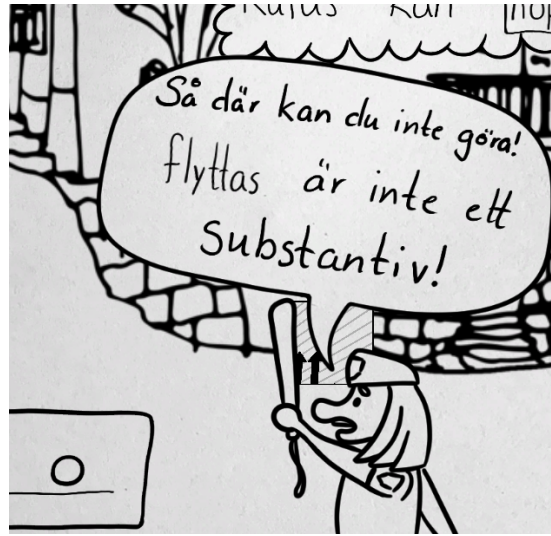
Figure 4.10: Screenshot from level 4 where the sentence "Spöke är en flyttas" is grammatically incorrect.

There where also feedback about the goal not being displayed clearly. One idea to better show the goal of the level was to add dialogue explaining it. However, having text telling the user where to go is restrictive (see section 2.2), and instead a guide in the form of a butterfly was added. This also reinforces another principle of level design (see section 2.2), since the butterfly clearly shows the objective of the level, but never how to reach it. For example, the butterfly might indicate that the goal is on a mountain, but there are several steps that the player needs to figure out on their own before being able to climb the mountain.

A new level based on *cooperation with an artist* (see section 4.8) was also introduced. However, this level was not of the same quality as the other ones due to time limitations.

4.10   Final Evaluation

The purpose of the final evaluation was to get a better understanding of how the user engagement was during the gameplay, as opposed to the two previous ones where the purpose was to further develop and iterate the prototype.

Step 1: Planning

The questionnaire in the final evaluation (see appendix A.1) was based on the revised user engagement scale (see subsection 2.3.2). There were 28 statements that the participant could agree or disagree with, using the Likert scale.

Step 2: Implementation

The final prototype was used for the final evaluation and it contained a total of seven levels. The first two worked as a tutorial, and the remaining five represented the three different design strategies.

Step 3: Testing

The participants played the first two introductory levels and was then asked to fill in some basic information, such as age and previous experience with puzzle games. They then played levels based on the different design strategies, these levels were divided into three sessions. After each session the player were asked to fill in a questionnaire based on their experience of that session. During the test the evaluator documented noteworthy actions taken and feelings expressed by the users, such as if the player got stuck, felt frustrated, or thought something was fun. During the time the participant filled in the answers in the questionnaire, the evaluator could ask follow-up question through an unstructured interview (see Figure 3.2.5) regarding the questionnaire answers or other noteworthy actions, this to get a better picture of the participants engagement.

Step 4: Data Analysis

The results gathered from the questionnaire and noted comments were all compiled and analysed separately from each other. The questionnaire answers were averaged out across the 21 participants and calculated. The qualitative data was analyzed similarly to the earlier user tests. Inductive thematic analysis was used to create themes to better understand the motivation behind the players' thoughts and actions.

5
Result

This chapter will present the results obtained from the final evaluation, which was conducted to measure user engagement while playing the game. Based on these results, the question "What is the impact of different design strategies on the user's engagement while playing?", will be answered.

The three level design strategies created during the development of the project will also be presented, and aspects to consider for future developers to think about will be described answer the second question of the purpose which was "What should be considered when applying different level design strategies?".

Also, one of the goals of this project was to develop a game to compare user engagement between different level design strategies, and this chapter will start by describing the final version of *AGCR*.

5.1   A Game Called Rufus

AGCR is a syntax-driven single-player puzzle game, where the player has to manipulate the game's rules in order complete the levels. This is done by dragging and dropping words into sentences displayed on the screen (see Figure 5.1.

Figure 5.1: Screenshot of the first level, which functions as a tutorial for the player.

When starting the game, the player arrives at the start menu. Here they can choose to start the game immediately, select level, or exit. If the player wants to select level, they are taken to a "World map" (see Figure 5.2).
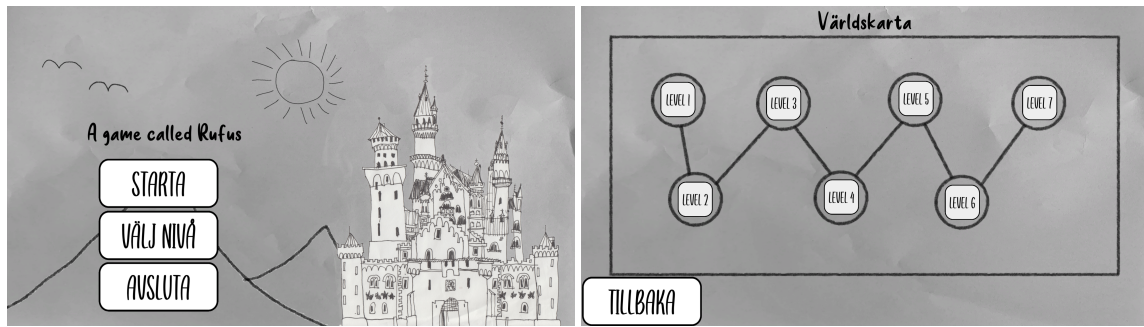
Figure 5.2: Left: Start screeen. Right: Level selection screen.

The first two levels are tutorials which familiarises the player with the mechanics and controls. Here, the player learns to manipulate sentences and moving the character. Level three and onward are designed using the design strategies and contain more puzzles.

The content of the game differs slightly from what was defined in the first MDA (see Figure 4.4). The major difference that can be observed in the final MDA (see Figure 5.3), is that the draw tool intended for character customization was excluded together with (NPCs), this as a result of time constraints.
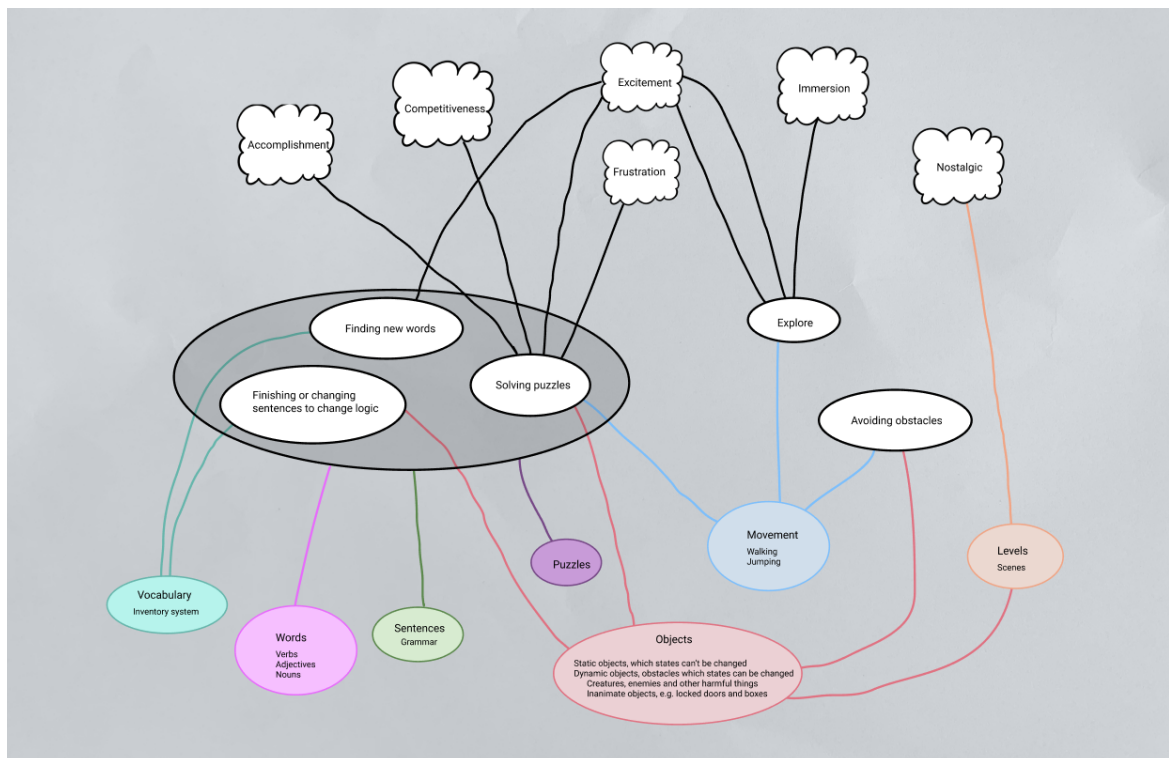
Figure 5.3: The final MDA.
Mechanics at the lower level, Dynamics in the middle and Aesthetics at the top level.

5.2 Obtaining the Results From Final Evaluation

This section presents the results from the quantitative questionnaire, as well as the qualitative data gathered. The questionnaire and material used for the final evaluation can be found in appendix A.1.

5.2.1 Quantitative Data

Looking at the results from the questionnaire (see Table 5.1 or appendix B.1 for all statements) and comparing the mean of the factors from the different sessions, there is one that stands out, namely *perceived usability*. While none of the means from the different sessions are above three (the neutral option), there is still a difference when comparing Session 3 to the earlier two, with Session 2 being ranked better.

Factor	Average value				
	1	2	3	4	5
Focused attention			3.2 / 3.4 / 3.4		
Perceived usability		2.0 / 1.8 / 2.6			
Aesthetics				3.9 / 4.0 / 4.1	
Satisfaction				4.1 / 4.1 / 4.1	
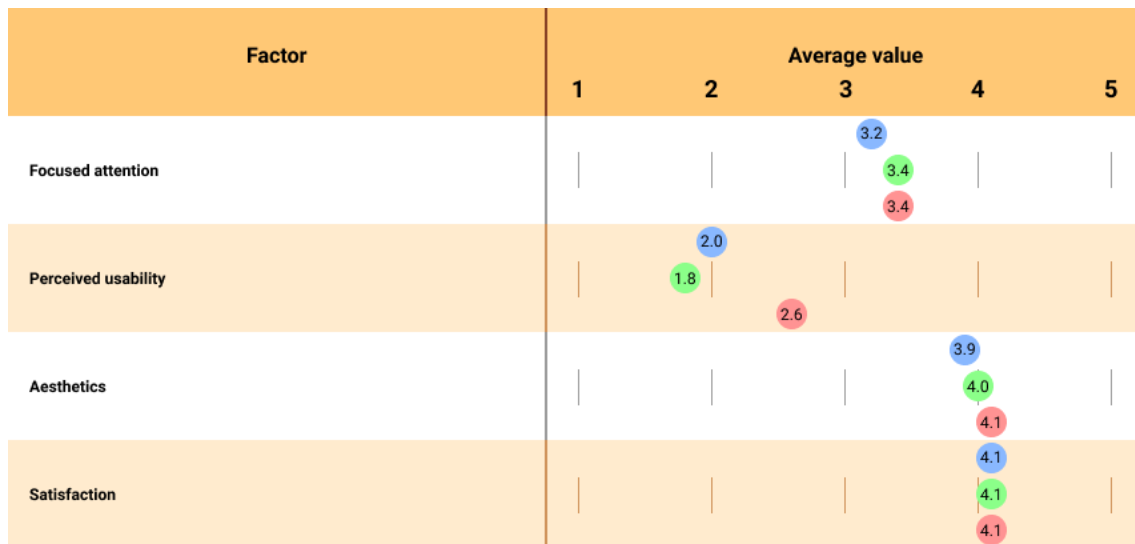
Table 5.1: Average of all values in the different factors. Average for Session 1 is marked with blue, the second green and the third red.

Other than that, there are some statements in focused attention that have quite different rankings among the three strategies. The statements are: *When I was playing the game, I lost track of the world around me*, *I lost myself in this gaming experience* and, *During this gaming experience I let myself go* (see Table 5.2). Interestingly, the first two statements' ranking increased in Session 3 while the last one decreased. This could be because of the greater challenge of Session 3, as seen in the higher ranking in perceived usability. The challenge might have immersed the players, while preventing them from relaxing.

Statement	Average value				
	1	2	3	4	5
When I was playing the game, I lost track of the world around me.		2.3	2.9 / 3.0		
I lost myself in this gaming experience.		2.7	3.1 / 3.3		
During this gaming experience I let myself go.			3.2	3.6 / 3.9	

Table 5.2: Average value for three specific statements in focused attention. Average for Session 1 is marked with blue, the second green and the third red.

5.2.2 Qualitative Data

The qualitative data was analyzed using the inductive thematic analyzing technique. Some themes could be recognized through all the different sessions while some were more specific to one or two. Note that all the following quotes have been translated from Swedish.

Aesthetic appeal was present during all sessions, but in a larger quantity in session 1. The aesthetics was something that most users enjoyed, and during Session 1, one person said:

> *"I thought that the game was pretty, it would have caught my eye if it was a real game".*

When reflecting on the experience as a whole, two others said:

> *"Very good. I thought that it felt like really good uniformly and nice, and it made the experience a lot more fun to play".*

> *"It feels fairytale-like, but still not extreme fantasy. It's something I imagine children would draw on paper if someone told them to draw a pirate ship or something."*

As seen in the quantitative data, this is something that was consistently ranked high among the users. There was only one user that commented that the illustration looked like they were drawn by a child, which was the intended sensation the player should get. *"Should feel like playing through a child's drawings in order to bring a nostalgic and playful feeling to the player."* (see subsection 3.1.4).

Enjoyment could be found through all sessions, although the reason for it was different for different users. Some found enjoyment from completing a challenging

33

task, for example one user said:

> *"When it was more difficult it was actually fun".*

but for others it was a source for frustration:

> *"I was frustrated by the trickiness in the game".*

and some preferred the less difficult levels:

> *"It was more fun now when it was easier, but the fourth one was also fun. But it was a little too hard too fast".*

Many users also expressed enjoyment from trying different combinations and playing with the mechanics. Looking back at the MDA in subsection 3.1.4, the feedback match the intended emotions that should be evoked the player while playing,*"Rewarding, both in learning purposes and enjoyment"*, as well as *"Sense of accomplishment after finally solving a tricky puzzle.".*

Difficulty and the variety throughout the sessions were expressed by multiple users. As mention in enjoyment, some liked the greater difficult found in the first and last session, while others preferred simple gameplay. Some users expressed that the reason for why Session 2 were easier than Session 1 could be because they had a better understanding of the game.

> *"Easier when one understood how the game was built, I had a better grasp of how it worked now. No major difference in difficulty compared to the last session".*

Other users implied that it could be because of the change in gameplay in Session 2. They felt that it was more of a focus on finding words, rather than solving tricky puzzles to progress:

> *"It was very straight forward, there was only one way. Find the words and move forward".*

As noted by almost all users, the last session was the most difficult one. Some thought that the jump in difficulty between the second and last session was too great and that it came very suddenly. Unique for the last session was that it was the only session where some users actually chose not to finish the level. This is also reflected in the quantitative questionnaire, where usability was perceived worse in the last session.

No clear distinction between game and background. In the first two sessions there were some comments about not being able to distinguish the game from the

background. For example one user pointed out:

> *"Sometimes it's a bit difficult to see what's there for aesthetic reasons in the background and what Rufus can interact with".*

after playing Session 1. A possible reason to why this was not expressed in the last session could be because the background was drawn with the level in mind, or as another user put it:

> *"There was not as much background which made it more clear where one could go" when referring to the more simplistic background."*

Disoriented was a theme based in the last two sessions. Comments like:

> *"I don't understand where I am.",* and
> *"Difficult to know where I am."*

were some comments expressing this. A reason this was not mentioned in Session 1 could be because low use of vertical puzzles. One user describes the problem in vertical puzzles as:

> *"I had to go up and down and I lost perspective of where I was and where I was heading".*

In the last sessions there were some comments regarding the combination of the teleportation doors and simplistic background:

> *"Yeah, like the graphic is of course simplistic, but it made it difficult to recognize where one is sometimes. In the beginning with the doors, because everything looked similar so it took some time before you got used to it".,* and
> *"You could see what you could jump on, but it was hard to orient oneself".*

As mentioned in the theme above, the simplistic background might make it easier to define game and background while making it harder to differentiate different sections of the level. Session 2 being difficult to navigate was not something reflected on in the quantitative study, as the statement *I felt that the game was confusing* (see appendix B.1) was ranked less than in Session 1. However, as mentioned earlier, Session 3 had a worse perceived usability and the statement was ranked much higher.

Lack of sound was only present in Session 1. A reason for this could be because the users were well aware of the lack of sound when playing the later sessions. While

some merely noted the lack of sound, one user expressed how the lack of sound took you out of the experience:

> *"You would be more engaged, or not more engaged but you would've lost yourself more in the game. Because right now I sit here and I hear sounds around me and it's stuff like keyboard presses and alike. You get distracted pretty easily".*

There were no items in the quantitative study that directly related to sound as the aesthetics focused more on the visual aspect.

Black and white were the only colors used with the exception of the butterfly that guides the player. Most users that commented on the lack of colors expressed that they would have preferred some color to the game.

> *"But I think that it was gloomy, it was dull because everything was so grey. Although that might've been the point".*

Target audience and comments about how the user did not feel like they were part of the intended audience were mostly clustered in Session 1, but could be found throughout. Many users thought that the game was developed for kids and that they would not recommend it to anyone in their own age group.

> *"It feels very simplistic and I think that it might be targeted to a younger audience because stuff like grammar with a or an ghost was shown. Sure, it's an error someone older can make but if feels like a grammatical correction aimed at the young".*

5.3 Level Design Strategies

One of the goals with the thesis was to create strategies to design levels that are based on pre-existing artwork. The three strategies are based on a general level design approach (see section 3.4), but differ in execution. In subsequent sections the strategies are described.

5.3.1 Level Design Based on Existing Environmental Assets

By designing levels based on assets already created, the designer tries to create different puzzles based on what is inside images. In this strategy, the background image decides how the puzzles should be formed and should therefore be chosen carefully. Choosing an asset with few details or noteworthy objects limits the type and amount of puzzles that can be created. Conversely, an image that is full of details or objects can make it hard for the player to identify what is background or

intractable objects. Since the images mostly contain environmental elements, there is a possibility that other assets have to be drawn to be able to create a puzzle. The idea is that even though some assets has to be created, the puzzle will mainly be based on the original asset. Of course, the added assets must fit the theme of the level and should not feel out of place.

Since this strategy heavily depends on what is in each background image, it can be hard to follow a checklist, but some general steps can still be created. Begin by looking at what can be objects with key functionality, meaning objects that can either help the player along with a puzzle or stop them from going a certain way. When key objects are found, a narrative needs to be established if not already done. The narrative should entail where the player needs to go and in which steps. After this is established, a regular reverse engineering process (see section 3.4) can take place. It can be advantageous to create a plan where all objects are marked out and important actions are highlighted, as seen in Figure 5.4.
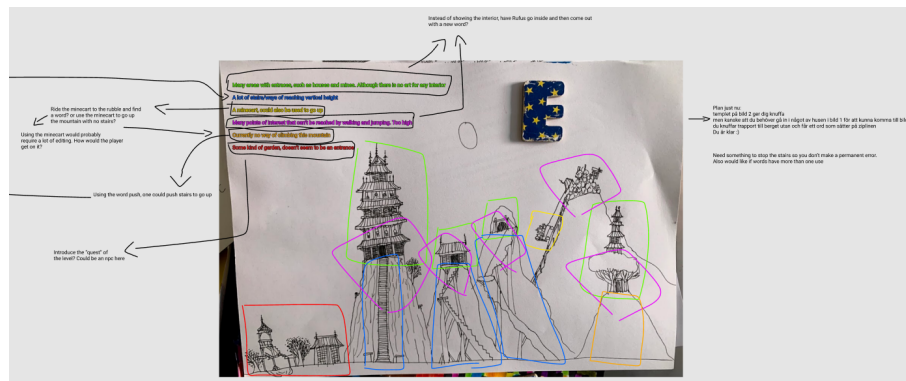
Figure 5.4: Early stages of level 6

The process can be divided into three steps:

1. What elements can be found in this picture?
Identify elements of importance in the picture and write down what makes them noteworthy. For example, they might be unreachable in the current state or they might spark ideas for a puzzle theme.

2. What can the elements be used for?
Go through the elements and define a goal for the player. Is it reachable? If it is, try to find obstacles. Are there any elements that can hinder the player from progressing? Or are there some that are needed to reach the end of the level? In that case, an obstacle could be to remove that element and reintroduce it later or create a different solution.

3. How can we get around or solve these obstacles?
If possible, use other existing elements as solutions. These could be elements that have been removed or moved in the earlier stage. If needed, a solution could also be

made from newly created objects not in the original drawing. As mentioned earlier, these should fit the theme of the level.

5.3.2   Creating Levels Through Puzzles

Using this strategy, puzzles are created and stored in a catalog to be used in future levels. A designer would then collect a few puzzles and connect them together to create one level. First, creating a puzzle can be done in multiple ways. For instance, a reverse engineering process can be used. The created puzzle would be designed with some characteristics or key aspects in mind, but would otherwise be general as to be easily used in any scenario. When a few puzzles are created, a designer can start to put them together starting with a core puzzle. If the level needs to be harder or longer, more puzzles can be introduced. After some puzzles are chosen and adapted to each other, the designer starts to look at assets to use to create the visuals of the level.

Since a level often consists of multiple puzzles that are intertwined, changing or creating a new puzzle can have unintended consequences. For this reason, the designer has to be particularly careful when introducing a new puzzle to a level in order to not interfere with other puzzles. This means that not all puzzles fit in all levels. However, this does not mean to discard the puzzle but for the current level another puzzle should be chosen.

5.3.3   Level Design Through Cooperation With an Artist

This strategy is very similar to how level design usually is implemented, meaning that levels and puzzles are created before the art. This strategy could easily be combined with other strategies but is especially good for attaining the same visual effect of the game. Since all of the images used in the level will be created by the same person, the art style will remain the same.

As for puzzle creating, this strategy is very lenient and makes it possible for the designer to create puzzles with a theme that is not already in the given assets, meaning that new obstacles and challenges can be created without breaking the theme. However, this strategy does not define any way of creating the puzzles, and the overall focus is rather on the theme and cohesion. This means that other strategies can be used together with this one.

5.4   Discoveries Using the Strategies

When designing a game, all levels should have a purpose and each level should have something new and/or exciting. To keep the levels from getting repetitive, it is important to change things up. This requires the designer to be familiar with multiple design strategies. This also ensures that the designer can pick the best strategy for

every new situation. Another way to ensure that each level is unique is by testing new ideas, and then conducting user tests to evaluate the level.

Sometimes when a certain strategy did not work, it was found that switching to another one could ease the design process. If inspiration was lacking, one could try to design the puzzles in a vacuum to utilize the modularity (see subsection 2.2.2) of such a puzzle. Designing a puzzle without taking where it will be used into account, can open up for ideas previously pushed aside due to environment restrictions.

However, only designing puzzles in a vacuum can lead to a hard time integrating them into the environment. For this reason, it was proved useful to design puzzles around objects seen in an asset instead. If this still does not work, the puzzle might still be good and can therefore be put in the puzzle catalog instead of discarded. Also, the artist can be asked to create an environment which fits the puzzle.

If a strategy does not work out, try to switch to another one. This will give another angle to combat the problems faced in the previous strategy. Furthermore, switching strategy as opposed to starting all over results in less work since achievements with the previous strategy can be reused.

It was also discovered in the final evaluation analysis that the strategies performed about the same in terms of engagement (see subsection 5.2.2). The data showed that *focus attention*, *aesthetics* and *satisfaction* stayed almost the same during all three sessions. *Perceived usability* was the only that fluctuated during the sessions.

5.4.1 Level Design Based on Existing Environmental Assets

Looking at the data from the questionnaire, this strategy preformed slightly better than the rest. Contradictory to that result, both users and the designers had their complaints with this strategy. Level 5 & 6 were created using this strategy and users felt that these levels were easier considering difficulty, according to the questionnaire and comments. Some thought this was a good thing, while others wanted more challenging levels. One reason to why this strategy was rated better in *perceived usability* can be because of the lower difficulty. In other aspects, the score was comparable to the other strategies.

Using this strategy was, from a level creators perspective, annoying and difficult. The levels always turned out too easy even though effort was put in to making them harder. Due to the lacking difficulty in the created prototypes, the level creation process was long and tedious. When a level was created and tested, it was described as an interactive experience rather than a puzzle. The levels were linear, meaning that completing a puzzle often did not contribute to the progression in other puzzles.

Furthermore, all of the assets used in this strategy were in some ways altered to fit puzzles or a particular context. Since the assets were already created, it was considered reasonable to edit the assets slightly. However, too much editing could

result in a product that strays too far from the original drawings. To allow for a bit more freedom, minor adjustments (such as raising ground and stretching empty space) were accepted. However, it was important that the original piece could still be seen and recognized in the final product, and the changes should only be done in order to accommodate for some mechanic or puzzle. As a designer, this needs to be defined from the beginning in order to be consequent since it will set the precedent for future levels.

5.4.2 Creating Levels Through Puzzles

Level 3 & 4 were created with the this strategy. The analysis, both from the questionnaire and from comments, showed that the users enjoyed these levels. However, it was not proved to be more enjoyable than other levels created with the other strategies.

From a designers perspective, using this strategy made it easy to come up with puzzles and then combining them. Creating puzzles in a "void" made it easy to let inspiration decide how the puzzle should be formed. An advantage of this strategy was that the puzzles does not need to be complicated, and it is only when they are put together they become complex. This means that puzzle creation is easy which is beneficial. However, combining puzzles together is more complex and takes more effort. This is due to the fact that just placing the puzzles one after another easily results in a linear experience. To create a more complex and interesting level, the puzzles need to be intertwined in some way. Not only does this create complex and interesting puzzles, it also prevents repetitiveness.

Repetitiveness is something the designer needs to be aware of when using this process. By using the same puzzles (even if slightly altered), the player soon figures out the solution and the difficulty is lost. To avoid repetitiveness, an old puzzle can be paired with new ones. However, using the same catalog of puzzles can get repetitive as well depending of the size of the catalog. One or a few new puzzles can instead be specifically designed for that level and paired with old puzzles from the catalog in order to break repetitiveness.

Because the sometime rigorous complexity and interconnectivity between puzzles that can occur, it is especially important that these levels are rigorously playtested. As a designer it can be hard to catch all bugs and unintended scenarios, and it is therefore beneficial to bring in outside personal that have never played the level and observe them playing. The designer should however keep in mind that playtesting the same level with a user multiple times will not be as revealing as having a new user play it. However, testing with new players is time consuming.

It was also found that if a level was too short, putting in a self-contained puzzle or additional content could extend it. The puzzle should be intertwined with the other puzzles in the level, for example by putting a key object behind the new puzzle.

5.4.3 Level Design Through Cooperation With an Artist

During the final evaluation it was noticed that level 7, which was created with the strategy (*cooperation with an artist*), was the most difficult one. The data from the questionnaire and the comments show that the majority of users thought this was the hardest level, and a few even gave up when playing it. Although this might seem bad, the focus attention of this level was the highest of all strategies.

Designing a level using *cooperation with an artist* was not as time consuming as the other strategies. However, this does not take waiting for the artist to create the assets into account. The design process of level 7 started out with designing the puzzles involved, this was done according to the reverse engineering principle. While designing the puzzles, the designer tried to think of an environment to the level that were shaped around/incorporated the puzzles. The designer has more freedom using this strategy than any other since the background is formed around the puzzles rather than the opposite. This also means that the designer needs to have a clearer picture of the level both visually and in regards to the puzzles.

When the puzzles were done and an idea of the environment were created, they were sent to the artist and explained. In our case Forsman created the requested assets based on his understanding and imagination. When using this strategy it is important to take into account that the artist needs time to create the assets. It is therefore suggested to work on other parts of the game while waiting.

6

Discussion

In this chapter, a discussion about the project will be presented. Thoughts regarding the final design of the game, the results from the final evaluation, as well as ethical and societal aspects will be discussed. The chapter ends with a discussion about suggested areas for further development.

6.1 The Level Design Strategies

When implementing the game the three strategies were used, this highlighted many of their strengths and weaknesses. Some of the strategies had disadvantages that other strategies did not. Realizing these, we quickly learned how to create levels with a certain strategy. Below are our experiences with the strategies shown:

Level design based on existing environmental assets The initial approach was to use a picture as a background and then create a puzzle that could match the picture. However, this seemed to block the creative process, coming up with something new that still fit into the picture turned out to be difficult. This led the level to be narrative rather than puzzle focused, as well as lacking a challenge. There could be many reasons behind this effect however we mainly think this is due to inexperience or having the wrong approach. This was solved using a more structured approach. Instead of looking at an image and trying to figure out a puzzle, we systematically examined the image and noted interesting features. This made it easier to implement the puzzles later rather than trying to do both at the same time.

Furthermore, it was decided that if something were to be changed it had to be minimal, such as making a closed door open or change the size of a tree. It took quite some time getting used to but once this way of thinking got more clear, it became easier to create levels that went along very well with the environment in the picture. An issue that occurred later in the process when there wasn't enough pictures left to create diverse and engaging puzzles using this strategy. The remaining pictures didn't have as many different elements than the ones that have already been used in levels before, which made it really hard to create something new and exciting.

Creating levels through puzzles Creating modular puzzles (see subsection 2.2.2) that can be reused and/or repurposed in other levels was an efficient way of creating puzzles. When creating puzzles they were saved to be applied in multiple scenarios. This lead to less time being spent on coming up with new puzzles. However, when

creating puzzles like this it was necessary to keep in mind that the puzzle should be able to slightly change depending on what assets or background is used, and what puzzles this level is composed of.

It became evident that puzzles should be self-contained. When creating a puzzle it should not modify other things outside of the puzzle or need input form the outside in order to completed. This is in order to minimize the risk of one puzzle interacting with another puzzle in an unintended way. This was often a difficult task and there usually had to be compromises, such as altering the puzzle or level slightly.

While using this strategy it was found that inspiration was important. Designing a puzzle in a "void" could sometimes be daunting, it was found that without any inspiration it was hard to get started on a new puzzle. When this was the case other puzzle games were looked upon for clever new puzzles to recreate or be influenced by.

Level design through cooperation with an artist This strategy is the most similar to the general level design process presented earlier (see subsection 2.2.1). Since there was no pre-existing artwork that was going to be used in the level, the level designer had free rein over the level. The designer had to take greater responsibility of how the finished environment should look but this also meant that puzzle ideas that did not fit any other environment could be incorporated anyway. Furthermore, this strategy does not define a way of creating puzzles, this is left to the designer to decide. After the project was done the group realised that this strategy is very similar to a regular level design process. That being starting with an idea, fleshing it out in a prototype and then lastly creating the art. The art for the level was drawn after the level and the puzzles in it had been designed.

According to the results, this strategy is worse when it comes to usability than the others. However, one reason for this result is believed to be the difficulty of the level and not the strategy itself. The qualitative data shows that the users felt that the last level were more challenging than other levels and this might decrease the perceived usability. Furthermore, since only one level was tested using the strategy, the results could easily be swayed by one bad level. To investigate this further more levels should be created and tested.

A major downside with this strategy was that the artist in this case was a child. Forsman could not be given specific deadlines, and if he did not want to draw the images for the game, there was nothing to be done. This resulted in long waiting times before receiving any results on the drawing for the level. The drawing used in the level was not in a finished state. This was because Forsman found it difficult to finish it. However, this situation was already predicted. To avoid similar situations in the future we believe that having more than one artist is good, and to set clear deadlines (if possible) when assets should be done.

6.2 Limitations

There are some limitations that are important to consider. Firstly, the final tests were done with quite few levels per strategy. Playing one or two levels might not be enough time to measure how engaged the player is while playing, since it is not enough time for the player to actually get engaged in the puzzles. One participant from the evaluation said "Too few levels to play before the questions about how one experienced the game. It was a bit to quick to play. Would rather have had longer sessions before answering questions about ones experience". Other participants had similar thoughts, especially about how short the first levels were. Another aspect of having few levels could be that poorly implemented levels affect the result greater when the sample size is small. Furthermore, the ideal would be to have several different versions of the game where all the levels are designed using the same strategy and contains the same goals. For instance, the levels should have the same goals and teach the player the same mechanics, as well as have the same increase in difficulty throughout the gameplay. It would have been even better to compare the level design strategies using different users to test each strategy. This to avoid the user getting affected by the strategies used in the previous levels.

Also, there is the fact that the testing were mostly done with friends and family due to the ongoing pandemic, this might affect the feedback in different ways. On a positive note, they might be more comfortable giving honest opinions and thoughts while playing. On the other hand, it is possible that the tendency for receiving positive feedback will be higher from people that you already know. Additionally the questionnaire is answered between the change of strategies during the final test, it is possible that the user remembers the questions and considers them further while playing the following levels.

The revised user engagement scale used to measure engagement was originally used on a gaming website. Since our intent was to measure the engagement of a game and not a website, some of the items have been changed slightly to better fit our needs. We have tried to keep the intent of the items the same to make sure that the same type of data is captured. We also translated the statements from English to Swedish. With all these changes in mind, it is possible that the original meaning behind the statements have been changed or lost.

Focused attention increased throughout the sessions, with session three being ranked highest. As mentioned in the results, session three was also deemed the most difficult one. A connection between difficulty and focused attention could be made, however session two had a higher score than the first in focused attention while being deemed, by many participants, as the easiest one. If session two's decrease in difficulty was because of the players' better understanding the game or simpler gameplay is not clear. Another explanation for the increase in focused attention could be that as they participants played they got more and more immersed in the experience as a whole. To combat these problems with measuring difficulty and focused attention one could have randomized the order of the levels. By doing this one can hope

to cancel the task sequence effect [36], that earlier tasks (or levels) affect how the participants experiences the next one. Although in *A Game Called Rufus*, early levels introduced mechanics to the player and therefore randomizing levels could not be used in this project.

6.3 Choice of Methods

This project regards game design with focus on designing levels, and can be defined as a wicked problem[28], since the solutions for this project can be either good or bad, rather than true or false. Using the method *design thinking*, as mentioned in section 3.2, is a solution-based approach that is effective when taking on wicked problems. Using this working method has facilitated the process throughout the project, and all members agree that this choice of method was successful.

Furthermore a revised user engagement scale (see subsection 2.3.2) has been used the to evaluate how successful the level design strategies was for engaging the player. This theory, on its own, did not give much insight to reasons for why user had certain feelings. However, together with qualitative data gathered, possible factors could be collected. The two gathering methods complemented each other well. This could be seen when aspects of game (e.g. aesthetics) consistent throughout the game was commented on. The qualitative data were perceived as higher in the first session even though the quantitative data showed no major difference between the strategies. This could be because users might comment on aspects more when they are first introduced or changed.

6.4 Societal and Ethical Considerations

Since the project results in a prototype targeted to a group including a younger audience it is important that the contents of the game are appropriate and correct. This was kept in mind through the entire process. However, something that turned out to be a problem was figuring out the appropriate difficulty level for the puzzles in the game. In order to make sure that most people (no matter age or other factors that might have an effect on the experience) feel included we designed the levels to be initially easy. This way, the player can progress through the game at their own rate. We do not want the player to get discouraged by a game that is too hard in the early levels. On the other hand, we also do not want the player to get annoyed or bored with too easy puzzles. To make the puzzle game credible and of reasonable difficulty, we evaluated all the levels by testing the ideas with players of all ages.

Another thing that was considered is that the drawings we use in the game are drawn by a child. What do we do if he is disappointed with the work or changes his mind about allowing us to use the drawings and withdraws his consent? In that case we decided to have a backup plan to make sure the project could be continued by replacing the drawings with new ones. We could make our own drawings to use as backgrounds in the game.

6.5 Further Development

This section will discuss different aspects which was not included in this project, mainly as a result of time constraints. Excluded game patterns, a higher number of levels as well as a game build for tablets can all be seen as important things to consider for further development.

6.5.1 Excluded Game Patterns

The majority of the game design patterns mentioned in section 2.1 were implemented in the final prototype, but there are some patterns that had to be excluded. This was due to to either lack of time, change of direction in development or simply that the game pattern did not fit into any of the designed puzzles.

The pattern *destructible object* got excluded in the design process since there was not enough levels to increase the difficulty of the puzzles to a level where a *destructible object* would be appropriate. Implementing such a mechanic would mean that the player might have to start over if they destroyed the wrong object. At that time, this could not be solved in a satisfying way.

Implementing *goal hierarchies* and using *scenes* was an idea at the beginning that also had to be cut from the final product since the direction was changed quite early in the process. From the beginning, the idea was that a level would be built up of several scenes, one scene consisting of a drawing, containing one puzzle. It became difficult to design since there was a limited amount of drawings and no time to wait to receive more. Eventually the design changed, using one or two drawings in a *level*, which affected the purpose of dividing a level into *scenes*.

One of the initial ideas was that the game should contain more stories, featuring a greater enemy, the witch "Suruf" that Rufus would clean up after in the game. However, *Storytelling* unfortunately had to be discarded due to lack of time. An narrative could have been a point of engagement for players by interesting them in the game [19].

Similarly, a draw tool for *character customization* was planned at the beginning. The idea was that the player could draw their own face on the player avatar and also name the character themselves. The idea of creating a draw tool for *character customization* was that it would let the player feel a personal projection, which is illustrated in our MDA (Figure 4.4).

6.5.2 Designing More Levels for Each Strategy

Another thing that would have improved both the gaming experience and the user tests is having more levels per level design strategy. In addition to this, implementing more levels would have made it reasonably easier to maintain the same difficulty for

the levels when comparing the level design strategies. There was not enough time to create as many levels as previously planned due to some delays and problems when implementing the mechanics in Unity.

6.5.3 Implementing the Game for Tablets

Moreover, implementing the game for tablets was an idea from the start that would have increased the accessibility for children to play the game, since it has been shown in a study by "Internetstiftelsen" in 2019 that for children tablets are more accessible than computers [37].

7
Conclusion

The purpose of this project was to answer the following questions:

> What is the impact of different design strategies on the user's engagement while playing?

> What should be considered when applying different level design strategies?

The data from the final evaluation showed that the strategies performed about the same in terms of engagement (see subsection 5.2.2). However, there were some interesting results regarding perceived usability for the level designed using the strategy "Level Design Through Cooperation With an Artist". Our theory is that the strategy used when designing this level, made it easier to create puzzles with higher degree of difficulty. This since there is no pictures from the start that restricts the layout during the creative process. Although how much impact the design strategies actually had on the user's engagement, could not be determined in this report and would need further research.

From a designer's perspective, when applying different level design strategies, it should be considered that in order to prevent levels from getting repetitive it is advantageous to change things up. To do this, the designer should be familiar with multiple design strategies. This ensures that the designer can pick the best strategy depending on the situation. Furthermore, testing new ideas by conducting user tests is good way to evaluate if a level works.

Insight into level design and user's engagement is offered, stemming from evaluations conducted with users. The first steps into developing novel level design strategies are taken, which can inform further research.

Bibliography

[1] Unity. Unity Technologies. [PC]. Accessed: February 11, 2021. [Online]. Available: https://unity.com/

[2] Epic Games. Unreal Engine. [PC]. Accessed: February 11, 2021. [Online]. Available: https://www.unrealengine.com/en-US/

[3] A. Teikari. (2019) Baba Is You. [Digital]. Finland.

[4] J. Roberts. (2017) Gorogoa. [Digital].

[5] S. Björk and J. Holopainen, *Patterns in Game Design.* Hingham, MA: Charles River Media, 2004.

[6] S. Björk. (2015, August) Clickability. Accessed: March 15, 2021. [Online]. Available: http://virt10.itu.chalmers.se/index.php/Clickability

[7] ——. (2011, September) Destructible Objects. Accessed: March 15, 2021. [Online]. Available: http://virt10.itu.chalmers.se/index.php/Destructible_Objects

[8] Nintendo, "The Legends of Zelda: Breath of the Wild," [Nintendo Switch], Japan, 2017.

[9] S. Björk. (2016, July) Inventories. Accessed: March 15, 2021. [Online]. Available: http://virt10.itu.chalmers.se/index.php/Inventories

[10] ——. (2016, June) Scenes. Accessed: March 15, 2021. [Online]. Available: http://virt10.itu.chalmers.se/index.php/Scenes

[11] ——. (2018, January) Non-Playable Characters. Accessed: March 15, 2021. [Online]. Available: http://virt10.itu.chalmers.se/index.php/Non-Player_Characters

[12] Blizzard Entertainment. (2021) World of Warcraft. [PC]. USA.

[13] Nintendo. (1985) Super Mario Bros. [Nintendo Entertainment System]. Japan.

[14] S. Björk. (2015, August) Winning by Ending Gameplay. Accessed: March 15, 2021. [Online]. Available: http://virt10.itu.chalmers.se/index.php/Winning_by_Ending_Gameplay

[15] A. Kramarzewski and E. De Nucci, *Practical Game Design: Learn the Art of Game Design Through Applicable Skills and Cutting-Edge Insights.* Birmingham, United Kingdom: Packt, 2018.

[16] GDC Advisory Board. Game Developers Conference. Accessed: April 12, 2021. [Online]. Available: https://www.gdconf.com/

[17] D. Taylor. (2018) Ten Principles for Good Level Design. Accessed: April 12, 2021. [Online]. Available: https://www.youtube.com/watch?v=iNEe3KhMvXM&ab_channel=GDC

[18] The Media Insider. (2020) How to read mise en scéne | Visual film analysis explained! Accessed: April 20, 2021. [Online]. Available: https://www.youtube.com/watch?v=ueSh66xktkk

[19] H. L. O'Brien and E. G. Toms, "What is user engagement? A conceptual framework for defining user engagement with technology," *Journal of the American Society for Information Science and Technology*, vol. 59, no. 6, pp. 938–955, 2008. [Online]. Available: https://doi.org/10.1002/asi.20801

[20] E. N. Wiebe, A. Lamb, M. Hardy, and D. Sharek, "Measuring engagement in video game-based environments: Investigation of the User Engagement Scale," *Computers in Human Behavior*, vol. 32, pp. 123–132, 2014. [Online]. Available: https://doi.org/10.1016/j.chb.2013.12.001

[21] H. L. O'Brien and E. G. Toms, "The development and evaluation of a survey to measure user engagement," *Journal of the American Society for Information Science and Technology*, vol. 61, no. 1, pp. 50–69, 2010. [Online]. Available: https://doi.org/10.1002/asi.21229

[22] Figma Inc. (2016) Figma. [Web application]. Accessed: February 11, 2021. [Online]. Available: https://www.figma.com/

[23] L. Torvalds. (2021) Git. Accessed: March 29, 2021. [Online]. Available: https://git-scm.com/

[24] R. Hunicke, M. LeBlanc, and R. Zubek. (2004, January) MDA: A Formal Approach to Game Design and Game Research. Accessed: February 11, 2021. [Online]. Available: http://www.cs.northwestern.edu/~hunicke/MDA.pdf

[25] Fireroot Studios. Design: greyboxing areas. Accessed: April 26, 2021. [Online]. Available: https://www.firerootstudios.com/design-greyboxing-areas/

[26] Jackw-gamedesign. (2016) What is Greyboxing? Accessed: April 26, 2021. [Online]. Available: https://jackw-gamedesign.tumblr.com/post/139960850160/what-is-greyboxing

[27] R. F. Dam and T. Y. Siang. (2021) 5 Stages in the Design Thinking Process. Accessed: April 8, 2021. [Online]. Available: https://www.interaction-design.org/literature/article/5-stages-in-the-design-thinking-process

[28] Interaction Design foundation. (2021) What are Wicked Problems? Accessed: April 23, 2021. [Online]. Available: https://www.interaction-design.org/literature/topics/wicked-problems

[29] H. W. J. Rittel and M. M. Webber, "Dilemmas in a general theory of planning," *Policy Sciences*, vol. 4, p. 155–169, 1973. [Online]. Available: https://doi.org/10.1007/BF01405730

[30] H. Sharp, Y. Rogers, and J. Preece, *Interaction Design: beyond human-computer interaction*, fifth edition ed. Indianapolis, IN: John Wiley Sons, Inc., 2019.

[31] J. Caulfield. (2019, September) How to do thematic analysis? Accessed: April 26, 2021. [Online]. Available: https://www.scribbr.com/methodology/thematic-analysis/

[32] C. Wilson, *Brainstorming and Beyond: A User-Centered Design Method*. Waltham, MA: Morgan Kaufmann, 2013.

[33] Interaction Design Foundation. (2021) Ideation. Accessed: April 8, 2021. [Online]. Available: https://www.interaction-design.org/literature/topics/ideation

[34] J. Rudd, K. Stern, and S. Isensee, "Low vs. High-Fidelity Prototyping Debate," *Interactions*, vol. 3, no. 1, p. 76–85, Jan. 1996. [Online]. Available: https://doi.org/10.1145/223500.223514

[35] M. Rosala. (2020, June) Rating Scales in UX Research: Likert or Semantic Differential? Accessed: April 21, 2021. [Online]. Available: https://www.nngroup.com/articles/rating-scales/

[36] A. Zippel. (2018, March) The Sequence of Tasks Makes a Big Difference in UX Testing. Accessed: May 10, 2021. [Online]. Available: https://blogs.perficient.com/2018/03/23/the-sequence-of-tasks-makes-a-big-difference-in-ux-testing/

[37] Internetstiftelsen, "Barnen och internet 2019," 2019. [Online]. Available: https://svenskarnaochinternet.se/rapporter/barnen-och-internet-2019/

A
Appendix

A.1   User Engagement Scale - Questions

Focused attention
- When I was playing the game, I lost track of the world around me.
- I blocked out things around me when I was playing the game.
- The time I spent playing the game just slipped away.
- I was absorbed in my gaming task.
- I was so involved in my gaming task that I lost track of time.
- During this gaming experience I let myself go.
- I lost myself in this gaming experience.
- I was really drawn into my gaming task.

Perceived usability
- I felt discouraged while playing the game.
- I felt annoyed while playing the game.
- Playing the game was mentally taxing.
- I found the game confusing.
- I felt frustrated while playing the game.
- I could not do some of the things I needed to do in the game.
- The gaming experience was demanding.
- This gaming experience did not work put the way I had planned.

Aesthetics
- I liked the graphics and images used in the game.
- The game appealed to my visual senses.
- The game was aesthetically appealing.
- The screen layout of the game was visually pleasing.
- The game was attractive.

Satisfaction
- The content of the game incited my curiosity.
- I would continue playing the game out of curiosity.
- I would recommend playing the game to my friends and family.
- Playing the game was worthwhile.
- I felt interested in my gaming task.
- My gaming experience was rewarding.
- The gaming experience was fun.

B
Appendix

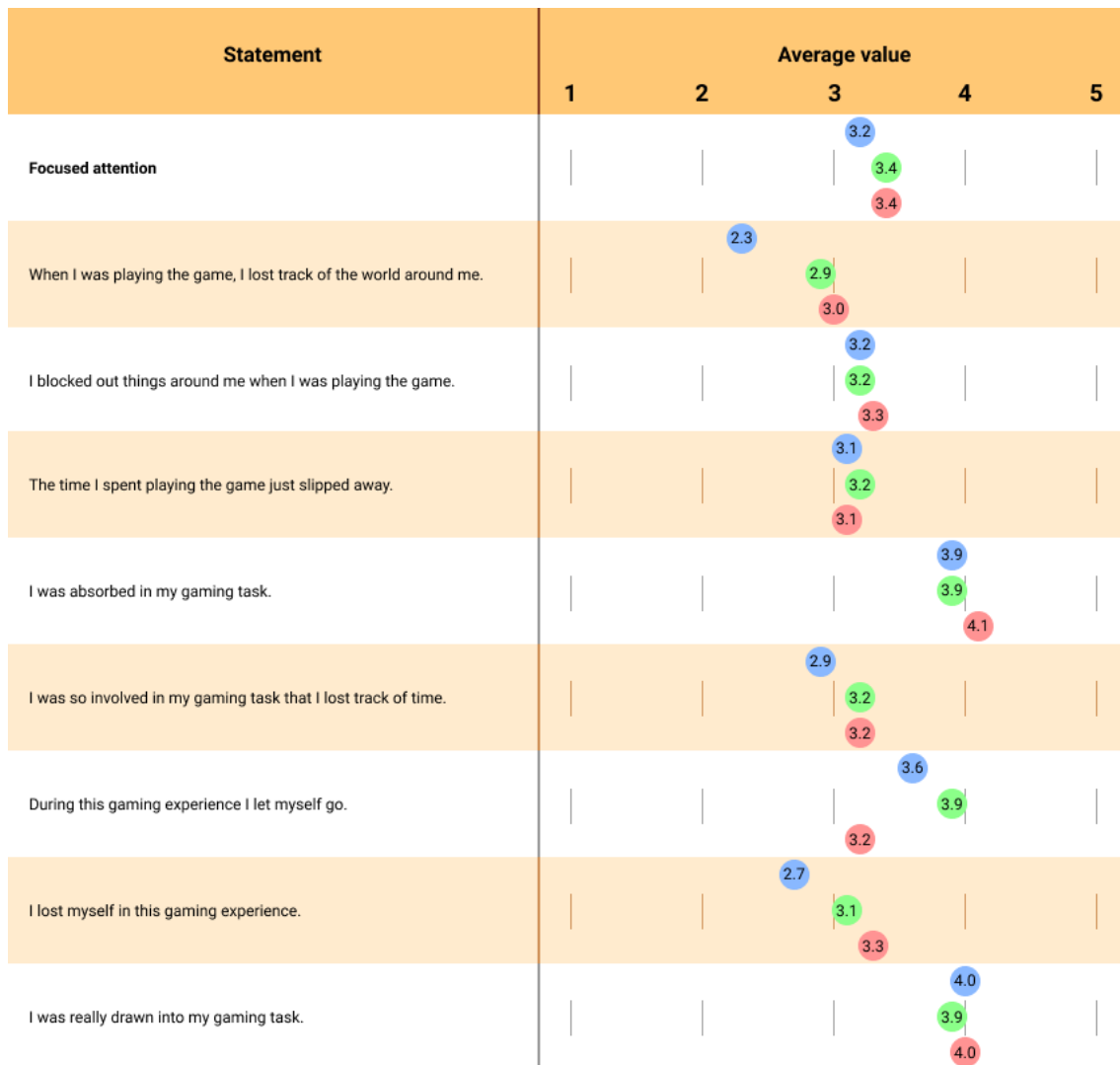## B.1 Values From Quantitative Evaluation

Figure B.1: Average values of the different statements in focused attention. Average for the first session is marked with blue, the second green and the third red.
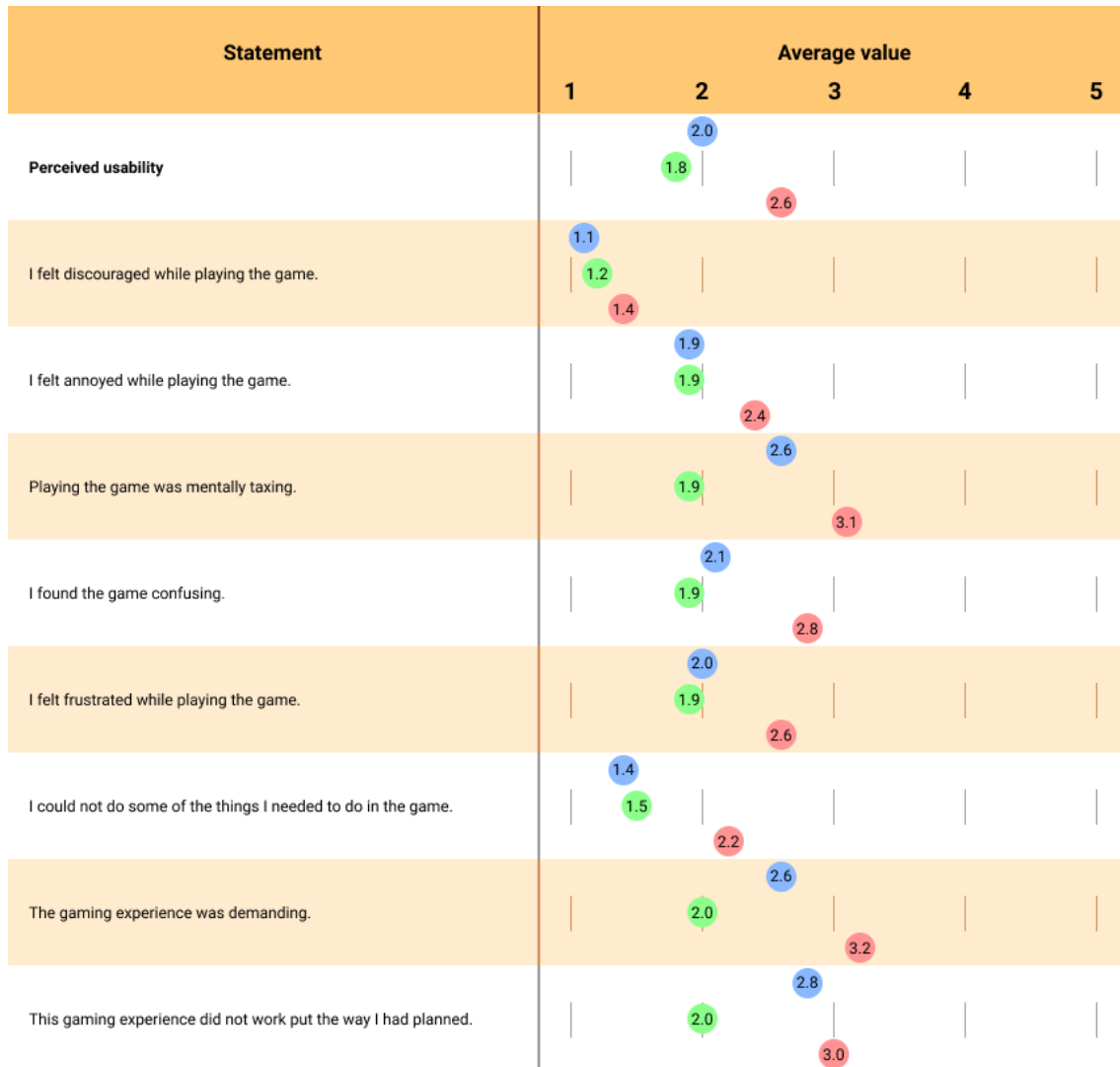
Figure B.2: Average values of the different statements in perceived usability. Average for the first session is marked with blue, the second green and the third red.
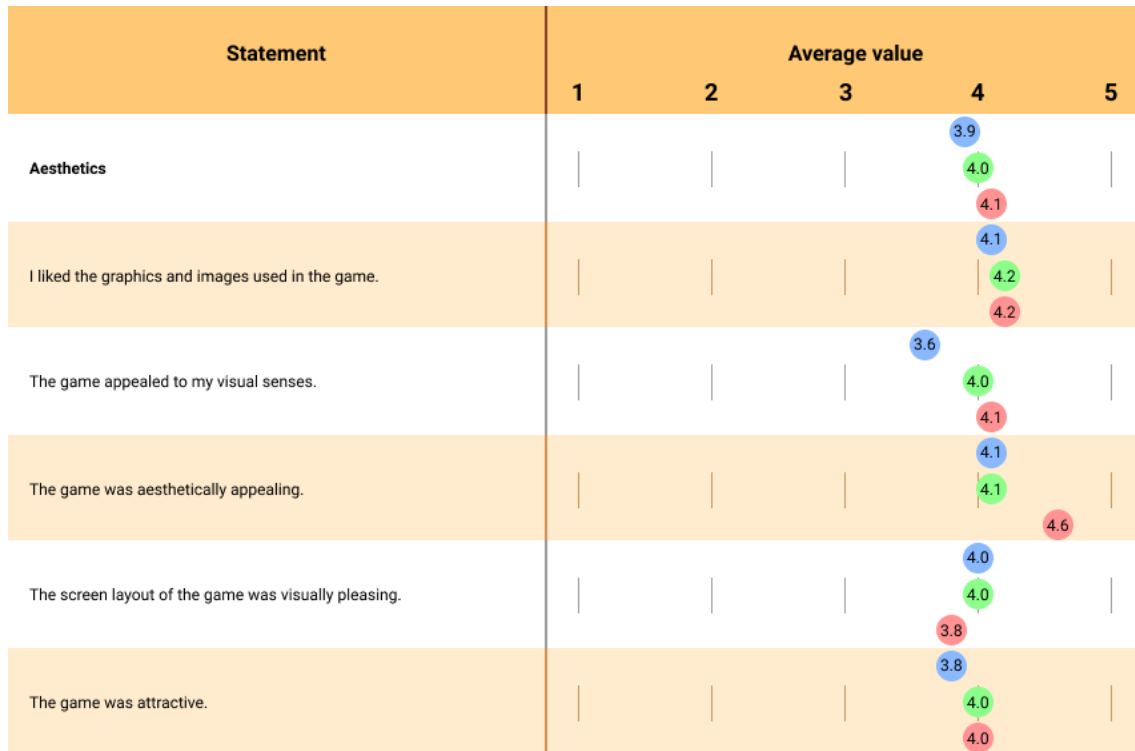
Figure B.3: Average values of the different statements in aesthetics. Average for the first session is marked with blue, the second green and the third red.

Figure B.4: Average values of the different statements in satisfaction. Average for the first session is marked with blue, the second green and the third red.