Joni Suojanen

Balancing and testing a skill-based progression system for a mobile card game

Bachelor's Thesis

Bachelor of Culture and Art, Game Design

Game Design

XAMK

**South-Eastern Finland
University of Applied Sciences**

ABSTRACT

The aim of the study was to figure out and test what is a good skill tree balance for commissioners' game. It was also researched what is considered to be a good pacing of skills and rewards in general.

Both qualitative and quantitative methods were used to research and test the concepts in the thesis. A contextual study was made to research known aspects affecting balance and motivation. Three testing methods were also used to find out game-specific behaviour in commissioners' game in the form of soft and hard data. Testing methods included physical interviews in the form of playtests, self-documentation in the form of a seven-day gameplay diary and game-specific analytics recorded from a specific period.

The study showed that balance can be game-specific, but there are many different routes to it. Theories such as self-determination theory and flow theory are used as concepts to understand and develop good experiences for the player and to seek a balance between elements of game design. The tests included in the study also pointed out that pacing the early skill tree and modifying its contents is essential for long-term playability in the game's current state, which has been and will be useful to the commissioner in the development of the game.

Keywords: game balance, game design, system design, skill tree, progression

CONTENTS

LIST OF FIGURES

APPENDICES

Appendix 1

LIST OF TERMS

KPI

Key Performance Indicators are metrics used to evaluate performance of a game in the form of collected data, such as churn or retention, users acquired and playtime (Kramarzewski & De Nucci 2018, 378).

Churn

A performance indicator that shows how many players stopped playing the game at a certain point or time. Churn is the opposite of retention, which depicts the number of players coming back after a specified time (Kramarzewski & De Nucci 2018, 383).

Playtest

Testing with a specified target group is done by players or a person employed to do it. Tests can be done with five to ten test participants focus-group. Playtests can also vary in length from minutes to several days. (Hodent 2023, 51.)

Roguelite

A game genre where the player must start the game from the start, losing some or all progression. In roguelites player typically has some permanent progression, such as keeping currency, experience or attained skills from previous runs. Roguelites derive from roguelikes, which have permanent death, without keeping any of the progression from the previous game. (Game Makers Toolkit 2019.)

Interview

Interviews are a qualitative research method, where an interviewer asks questions to get data. Interviews are typically conducted with one or more people per session at a chosen location. Interviews allow good control of the test situation and can be used to collect individual opinions, preferences, and experiences. (Cote & Raz 2015.)

1 INTRODUCTION

Game balance is an important part of the game development process, which can affect games success in the long term. Balancing the game system and its economy is vital for a good player experience. An unbalanced game can in the long-term drive players away and increase game churn rates. The target audience must be understood to allow variety, choice, and fun for the player. A good gameplay experience allows players to stay in the flow and be motivated to play the game one way or another.

The primary research question for the thesis is "What should be considered balancing skill tree in *Break the Beyond -game*? To answer the question both quantitative and qualitative testing methods are used. The author's initial hypothesis for the primary research question is having proper timing and well-paced rewarding feeling while playing the game. Players should be able to progress at the start to hook them to commit to the skill tree and allow motivation to play the game.

This thesis aims to investigate a suitable balance for the skill tree in the commissioner's game and to test the skill tree by using both quantitative and qualitative testing methods. The skill tree is connected to player's progression in the game. Therefore, the thesis also aims to answer a secondary research question: "What pacing of player's progression is considered good?" As the question can have varying answers and is contextual, the author aims to answer the question from the perspective of the *Break the Beyond* -game. To address the question above, tests are conducted to track the player progression over time, while following players' motivation to continue. The hypothesis is that players receive the first two skills within 30 minutes of gameplay to hook them to stay longer.

This thesis includes balancing and testing a mobile game project: *Break the Beyond* by Latchback Games. The game was launched for public testing in the summer of 2023 and is available for download in the Google Play Store for free. *Break the Beyond* is a mobile card game, with merging and roguelite elements.

Physical playtests are also performed as interviews with an aim for 20-40 test participants per testing location. The tests will be performed in two locations with separate game builds and participants in the form of an AB test. They are performed to gain insight into players' progression and motivations for playing the game. The thesis will also utilize analytics received from open testing, which allows a higher number of participants and quantitative data gathering. The testing will be performed online using a game build in *Google Play* and *GameAnalytics* service. The online test will be performed in the form of a gameplay diary using user participation using self-documentation. Feedback from the game will also be gathered through multiple sources: Games community discord, google forms and feedback from physical interviews.

As the thesis will cover an in-development game, the thesis' quality and reliability may be affected, due to possible production-related risks, such as the project getting postponed or ended prematurely, non-disclosure agreements or delays in the development process itself. As the research is codependent on the commissioners' development schedule, everything related to the development of *Break the Beyond* can affect the possibility space of the thesis.

The thesis started as research on how the balance of the game affects the game's player retention. However, due to uncertainty regarding the ability to reach statistically significant results, the subject of the thesis was changed. In addition to this, the system to be balanced was restricted to the skill tree system, as balancing the whole game in the early stage of development is both difficult and a waste of a small team's resources.

The thesis covers the areas mentioned above by starting with the introduction of methods below (Chapter 3). The chapter will start with methods used to answer the thesis's primary and secondary research questions. This chapter will also contain more detailed information on the timetable and Latchback Games, who is the commissioner of the thesis. The author recommends referring to the list of terms above this section of the thesis to understand the terms used in some of the following chapters.

The Chapter 4 introduces the contextual study regarding the balancing theory researched in the thesis. The study covers common dials to keep in mind when balancing a mechanic and tools to implement the balance. The chapter will also cover theory regarding modern testing methods relevant to the thesis to give a better overview of testing in games in general. Chapter 5 will cover what is a skill tree, and what are common ways of implementing it concerning balance. The chapter will also contain skill trees meaning in the game system.

The Chapter 6 contains more detailed information regarding the performed tests that are included in this thesis. The test plan for the thesis is broken down, containing more details about performing the tests individually. The chapter ends with an analysis of the results. The thesis ends in the conclusion, where the thesis will be broken down to its main points, presenting the findings and hopefully an answer to the research questions. Regardless of the results a post-mortem is also included in the conclusion.

2 COMMISSIONER

The results of the research will be utilized as part of the commissioner's game in development: *Break the Beyond*. Latchback Games was founded in 2022 by a group of friends from Lahti University of Applied Sciences. During the time of their founding, Latchback Games was developing its pilot project: Crankshaft Crust. Currently, Latchback Games develops the game *Break the Beyond* that the thesis focuses on. *Break the Beyond* is a mobile card game with roguelite and merging elements. The game also has adapted RPG elements to the game, aiming for it to be an easy gateway to the genre.

The game is about a wanderer in a fantasy-based world that is being consumed by an unknown entity. The realms are slowly changing, which is noticed by the characters, and thus they try to find out why. Players must defeat enemies by merging weapons from looted items, while unlocking new skills to discover new abilities and progress further (Figure 1).

Figure 1. The core loop of *Break the Beyond*
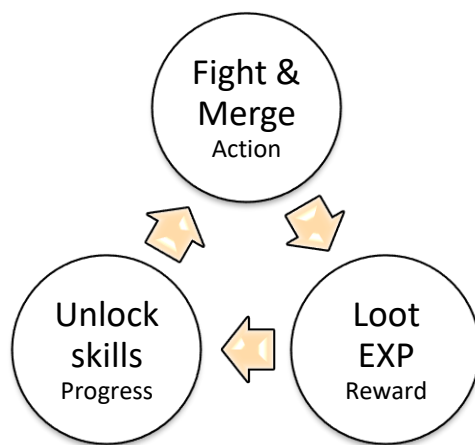
For *Break the Beyond* (Figure 2), the intended target audience is casual and mid-core players. This must be considered, as developing the game accessible to newcomers is important. Casual players might not be very familiar with the fantasy-themed role-playing game mechanics or the roguelite element the game is aiming to achieve, which means the game has been relatively easy to get into.
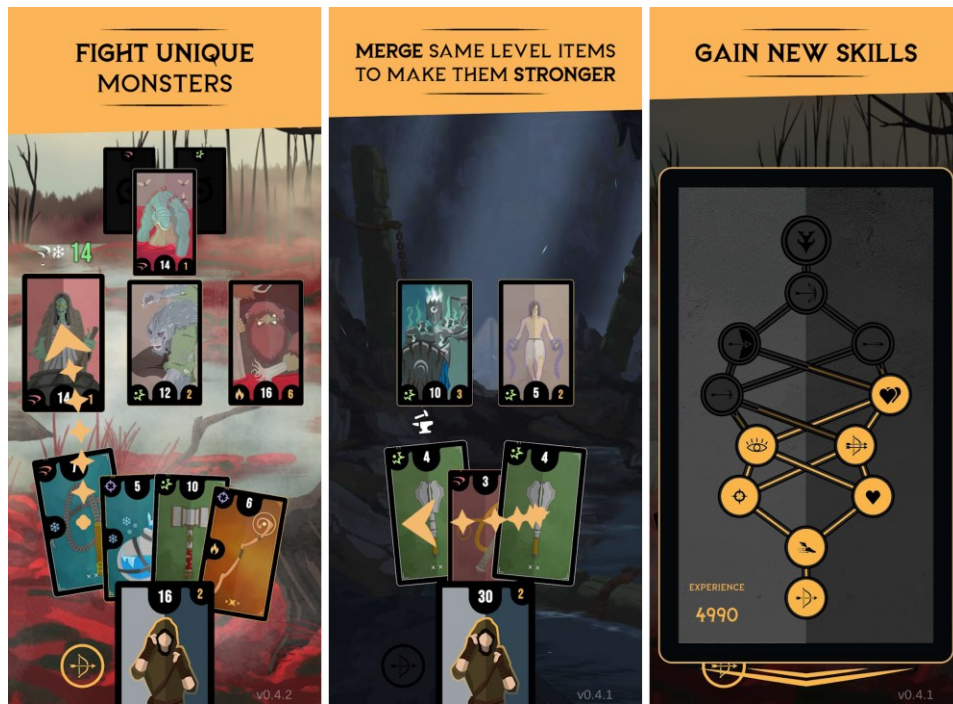
Figure 2. *Break the Beyond* (Latchback Games 2023)

In the early results developers could see that many test participants liked the core game and described it as addicting. However, the game's retention has not been as high as should. This was hypothesised as a lack of goal and reason to come back, meaning many of the players do not have a motive to return the game. This is why a skill-based progression system was developed for the game.

3 RESEARCH DESIGN

To answer the primary research question: What should be considered balancing a skill tree in the *Break the Beyond* -game? Theory regarding game balance is used to create a contextual study, that gives perspective on common methods to affect the balance of a mechanic.

3.1 Framework and Research Question

The quantitative and qualitative research methods are used to answer the secondary research question: What pacing of players' progression is considered good? To answer the question an AB test is performed as an interview at two locations: 1) South-Eastern Finland University of Applied Sciences and 2) Lahti University of Applied Sciences. Soft data is also

gathered at the same time from a separate testing session, which is performed online utilizing probes, which allows more specific qualitative data in the form of a gameplay diary. In addition to this, quantitative data is gathered from the commissioner's game to reference the results from soft data and hard data together to form a better understanding of the subject.

In Figure 3. player motivation is presented as the core for balancing progression. The thesis utilises quantitative and qualitative testing methods and try to answer the research question regarding players' progression and its relevance to its reward. As the overall balance would be too large a scope, the author limits the balancing to *Break the Beyond*s skill-based progression system, also known as the skill tree.
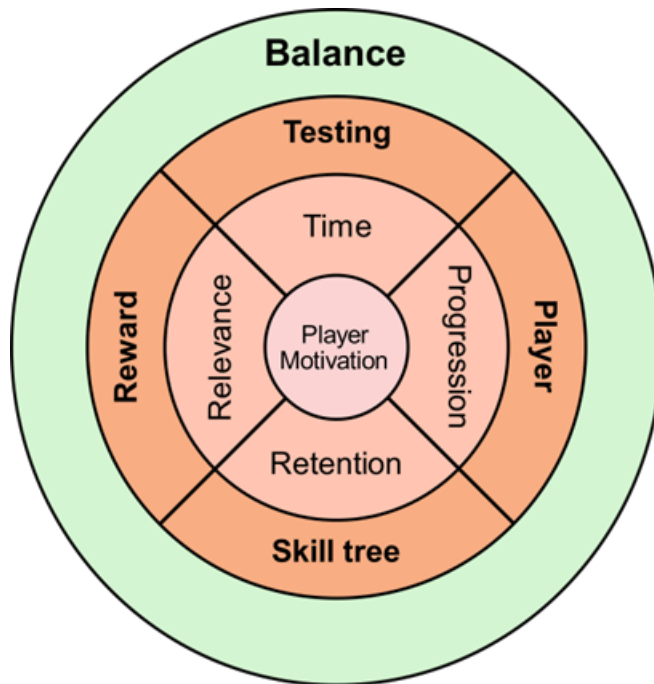
Figure 3. Framework of thesis

The thesis starts with a contextual study of researched sources regarding balance, progression, and motivation. The section strives to answer how balance is perceived, and what there is to consider in order to balance the commissioner's game, which also creates a base for the primary research question. After the theory regarding motivation, skill tree is covered as a progression system. As the thesis focuses on skill tree, it is covered in more detail compared to other progression loops, outside of the scope.

3.2 Research Methods

Afterwards, the thesis will cover theory regarding testing methods, such as playtesting, gameplay diary and analytics. These are all valid ways to test games in general. The tests are performed and analysed at the South-Eastern Finland University of Applied Science (XAMK) and Lahti University of Applied Science (LAB). The tests try to answer the secondary research question. The thesis ends with a conclusion, which goes through the main points of the thesis and results. To gather different viewpoints to the research question multiple methods are used as presented in Table 1.

Table 1. Research methods of thesis

Method	Data	Participants	What
Playtests + Interviews	Qualitative	XAMK (14) LAB (4) From other (2)	• What is the player's experience and preference for using the skill tree in one play session?
Gameplay Diary	Qualitative	Discord (9)	• What is the player's experience and preference for using the skill tree in a week?
Game Analytics	Quantitative	Players (80)	• What is the player's behaviour in unlocking skills and how much experience does the player unlock according to hard data?

These locations were picked due to being easier to access and gather participants. Interviews have the benefit of being more casual compared to other methods, performing tests in a controlled environment is good for achieving more accurate results. As the test participants are required to play the game, it is more beneficial for test participants to have the possibility to enter a state of flow to allow more genuine results in the process. (Cote & Raz 2015, 106.) It is also more beneficial for the author to see participants' facial impressions and body language while playing, which makes an interview a suitable research method (Hirsjärvi et al. 2009, 205).

While the main research is conducted on students and faculty of previously mentioned locations, some additional people were randomly picked from industry-related events. In addition to physical interviews, a test utilizing probes was conducted in the form of a gameplay diary. These methods are

meant to open insights into test participants' daily factors in human lives (Mattelmäki 2006, 42). In this case player's progression, experience and motivation will be measured for the seven days utilizing the diary analytics gathered during the time of testing the game.

Participation in the tests was not limited in any way and was conducted at each location separately at arranged times. The author used a booking system Koalendar to arrange times for participants over the weeks for better management and easier participation. A gameplay diary was conducted with ten participants, from which each filled out a seven-day diary using Google Forms. The details regarding the gameplay diary are discussed more in Chapter 6.

All registered participants were invited to select a suitable time to perform the tests. Participation was advertised at the previously mentioned locations using fliers and pop-up booths for the test. In addition to this, the participation was advertised on social platforms such as discord communities of said locations. Latchback Game's squeezable rubber duck-stress toy was also added as a lottery prize to motivate the participants to join. Test participants were also invited to join testing with gameplay diary as an addition. No reward was offered for participation in the diary. The total amount of participants was divided into A and B to test different options for the skill tree. Each participant was asked a question when booking a time:" What kind of games do you usually play?" and "Have you played the game before?". These questions gave insight into their overall experience in games, which affected the results of the test (Table 2).

Table 2. Research methods pros and cons

Method	Pros	Cons
Playtests + Interview	• **Allows perspective into what is players' behaviour short term and why.** • Experience and preference gathering (Cote & Raz 2015). • *More relaxed situations* • Controllable situations.	• Possibility for bias. • Ethical considerations such as privacy policy, research permissions and proper conduct of tests.
Gameplay Diary	• **What and why is the players' behaviour in the long term?** • Easy to perform. • Can be performed Online. • Can open other questions. • Data in a natural environment.	• Less control. • Needs planning. • Platform restrictions.
Game Analytics	• **Hard Data** • Requires third-party SDK integration. • High amounts of data. • Dependent on marketing. • 24 / 7 data.	• Easily misunderstood. • Requires developers. • Less control.

During the analysis, possible existing and upcoming problems were searched to implement better product in the future. Each case was analysed separately and anonymized if needed. After the analysis of results, the reward for test participants was picked randomly. The recipient of the reward was contacted with an email to arrange a suitable delivery. The validity of the results could have been affected due to many reasons. As the overall number of participants was low, it could not reach statistical accuracy. As the sample group was not necessarily part of a specified group, the results can vary greatly. Such things as players' experience, preferred games and past playtime in the game could affect the test results. Interviews have a chance of having bias if conducted poorly or interrupted. It should be noted however, that research gathers opinions and resource usage.

To consider the ethical aspect of the study, each location has approved author's research permit requests. A formal test plan has been filled with test cases, considering biases the test performer or test participants might have. Participants have also been informed regarding any information of the study

both in text and interview. Participants were informed that they could withdraw from the research at any time before and after the test session by contacting the author. As the testable product will be a prototype build of the game, it can be prone to bugs that have a different experience compared to the released product.

3.3 Timetable

The research started in June 2023 with subject creation and its acceptance by the commissioner. During the preplanning section of the thesis, its scope was decided, and the initial process was ideated and planned. After preplanning the main research was conducted referring to the subjects' main references and sending research permits to conduct accurate and ethically correct testing sessions.

As presented in Figure 4, the preplanning and research periods were performed during the summertime. The communication with test participants started during the implementation phase in August 2023, which included writing the acquired research in the form of a contextual study. The information was contained in the contextual study after being used to improve already planned testing sessions, and implementation of the thesis process.

JUNE 2023　　　　　　　　　　　**SEPTEMBER 2023**　　　　　　　　　**DECEMBER 2023**

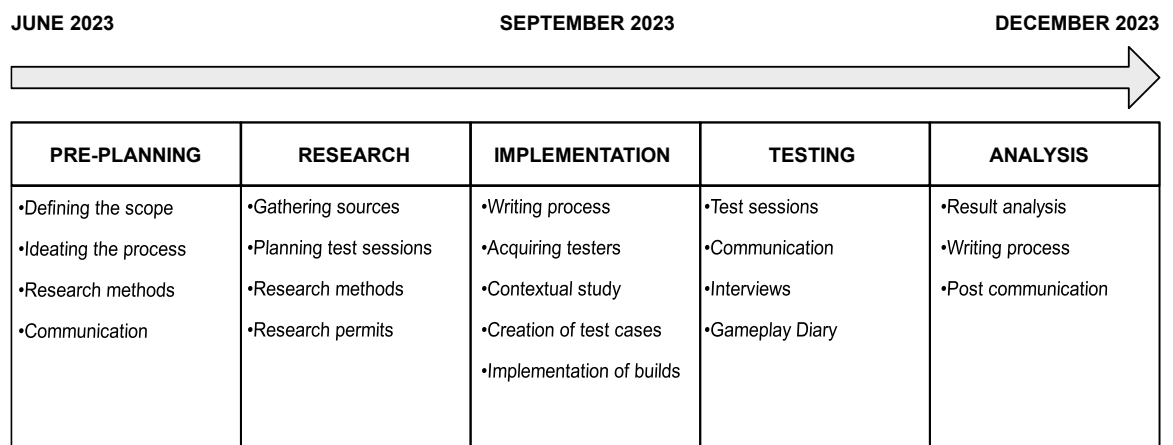| PRE-PLANNING	RESEARCH	IMPLEMENTATION	TESTING	ANALYSIS
•Defining the scope	•Gathering sources	•Writing process	•Test sessions	•Result analysis
•Ideating the process	•Planning test sessions	•Acquiring testers	•Communication	•Writing process
•Research methods	•Research methods	•Contextual study	•Interviews	•Post communication
•Communication	•Research permits	•Creation of test cases	•Gameplay Diary	
		•Implementation of builds		

Figure 4. Thesis timetable

The acquirement of test participants for the future was started using the Discord platform after the initial implementation period. Direct communication towards the target audience was established. After acquiring the required test participants for this research, the tests are performed according to the test

plan at two physical locations. The tests were performed as an interview, following another session with the gameplay diary as the main method of research. Communication between included parties was done between these two phases until the end of the research. The results of the test were then analysed and documented. The results are then communicated to the required parties.

4 GAME BALANCE AND PLAYER MOTIVATION

This section of the thesis strives to introduce what is considered good game balance, what it consists of, and what are common practices according to information gathered from research regarding the subject matter. According to Schell (2021, 216), the time to balance a game should be half of the production time and is considered a very difficult task in itself. Balancing a game should start when a game is playable, however, the timetable can shift between larger and smaller studios.

4.1 Dials of Balance

Game balance has many different forms and can be perceived with a variety of opinions. A skilful player might argue a game to be easy, while a new player might consider the game to be difficult. The game can also seem balanced according to numbers in designers' sheets but is very unentertaining or not fun to play. The game flow must be engaging to the point the player wants to continue the game. There are many ways that a game is unbalanced.

According to veteran game developer Schreiber and game director Romero (2022, 10.), three measuring dials make the game fun: difficulty, quantity, and timing. Game can be too hard or just right, which can be interpreted very differently depending on player skill, and expectations. There can be too many resources or components, such as enemy health can be too high compared to the damage the player can inflict. Opposite to the software industry, games tend to maintain a perfect balance in game flow, which requires adjusting many dials of design, iteration, and testing.

The designer's intent, the target audience and the purpose of the project are also considered when adjusting the dials mentioned above. A game meant for the casual target group can be too easy for a hardcore player with lots of experience in games in general and vice versa. A designer can also have a planned intent for the target audience, which affects the shape of the game. It is also important to consider the purpose of the project: an artistic game might not try to make the game balanced, but rather an experience, while a mobile free-to-play game's purpose is to gain revenue for the company that developed it. (ibid. 10-11)

Difficulty is an effect of physical or mental strain. A person with a limb impairment could find interacting with a game more difficult than a player with a good sense of games in general. (Schreiber & Romero 2022, 39-40.)
The difficulty is perceived to be balanced when a player is in a state of flow. Flow is a player's state where they are focused to the point, where the outside world becomes unnoticed. When the player enjoys the game to the fullest and is fully immersed in the gameplay. However, the player's state of flow can be interrupted if gameplay causes the player to be anxious or bored. The player can also not experience the flow state if the initial task itself is not understood or the overall focus cannot be achieved, as the task does not exist or is unclear. (Schell 2021, 144-148.)

According to psychologist Csikszentmihalyi (2014, 115), who is the founding father of flow theory, the state of flow is most found in the workplace or a task the person enjoys. The flow state does not occur in leisure time, as there are no clear tasks planned due to the time being allocated for relaxing. Making the game too difficult disrupts the flow and can be so frustrating that the player will move on to another game. There is no global rule that works for all games, however, games can interpret other similar games, in order to have similar experiences.

Every game balance consists of values and numbers that are part of its economy. According to Schell (2020, 213), there are differences in types of games. If the players start with different values, it is called an asymmetric game. If players start with the same values, the games are considered to be symmetric. The difference between asymmetric and symmetric games is in

the fairness and possibilities of the player. Players have same possible situation if they have an equal starting point. An asymmetric game has varying starting points, which will affect the possibilities of each player. Sometimes this can also increase the probability of winning for the other player, meaning they have an advantage over the other.

Roguelikes and roguelites tend to use randomization to increase both the replayability and length of the game. The loss of progression or some of it is close to the core of the genre, which requires high replayability and an engaging core that has been balanced to optimal player experience. However, randomization can be detrimental to a game. According to game designers Kramarzevski and De Nucci (2018, 53) randomization can risk the understandability of game rules and cause uncontrolled extreme situations between too-hard and too-easy sessions. Negative effects are usually more noticeable than positive effects, which makes unbalanced randomization a risk to long-term gameplay in general.

Taking all things into consideration requires a mindset of better user experience. Balancing a game is a struggle of making countless informed decisions that can cause conflicts and trade-offs in another situation. (Hodent 2021, 24-29.) Users can experience many things from mentioned above to what is engaging or "fun", however, fun cannot be used as a measurement dial as it can be subjective to opinion (Hodent 2018,135-136). Players form opinions based on their experiences, which is the main reason balancing a game is difficult. A specific target group for the game must be made, as the balancing dials are made with its target users in mind. (Jaffe 2013, 4.)

By creating experiences for the target audience, the designer has a better grasp on where the limit between boredom and fun will feel to the player. A fun game will keep players, but a boring game will send players away due to a lack of stimulation to the brain. (Koster 2013, 42.) Fun however is a broad measuring tool, which some researchers prefer to avoid. Instead, many refer to flow theory, as it uses engagement and frustration as a measuring tool. Technically designers could measure fun and boredom in the same way as engagement and frustration in flow theory, but measurement does not necessarily allow any insight into why something is fun or boring. It could be

argued that when the player feels fun or bored, something is done right, however, fun might also just mean the player enjoys receiving strong weapons in the game.

4.2 Player Motivation and Rewards

Player motivation is important for sustainable player engagement regardless of how it is measured. Without motivating the player, a game can easily become both frustrating and boring. The game's retention or churn rates can vary greatly if the player is not motivated to continue playing the game.

At the time of writing the thesis, the main motivation for the player in *Break the Beyond* is to gain experience and unlock new skills to perform better in the game. The longer the player plays, the more content they can discover with new unlocked skills. The skills, however, have not been tested and paced any more than in theory. The proper timing of gameplay and careful pacing of rewards are an important part of balancing a system. A designer can adjust the pacing of pressure, stress, and rewards to affect the game flow (Hodent 2021, 35). The player can feel betrayed if they are not being rewarded enough for the time they have spent on the game, causing resentment that is hard to unravel. The pacing of the game can also be too fast or slow, causing the game to be too intense or too slow for an average player. (Schreiber & Romero 2021, 56.) There is also a relation to how gameplay is timed on different genres, as designers might try to replicate the length of a successful game (ibid. 469).

There are many ways to pace rewards for the player that can be split into based on time and actions – (1) interval rewards based on time passed and (2) ratio based on actions performed by the player – or the probability of the reward that is continuous and intermittent as shown in Figure 5 (Zubek 2020,138; Hodent 2018, 63–64).
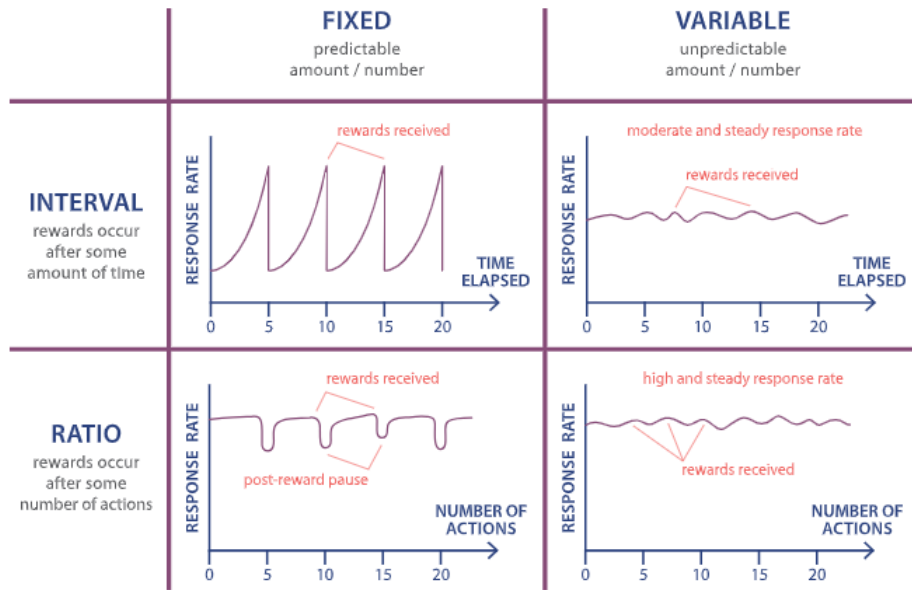
Figure 5. Different types of intermittent rewards and their approximate impact on behaviour (Hodent 2018)

Zubek (2020,138) on the other hand describes many variants of rewarding schedules as a way to plan out the pacing of rewards as seen in Table 3 with example use cases from *the Break the Beyond* game.

Table 3. Different rewarding intervals with examples from *Break the Beyond* (Zubek 2020)

Reward type	Meaning	Example case
Continuous	Rewards are given directly for each action	EXP given per defeated enemy
Fixed interval	Rewards every *n* second when performing actions	Daily XP boost
Fixed ratio	Rewards every *x* action	XP to unlock a new skill
Variable interval	Rewards at randomized points in time when performing actions	Random loot drops rewards received based on the probability factor
Variable ratio	Rewards every randomized number of actions	Special relics sometimes received from treasure chests and defeating a boss

Skills are directly relative to the amount of time and effort the players use in the game. Player receives skills over time, instead of all at once. By doing this a hook can be created to motivate players to continue a little longer at a time. Motivation can be divided into multiple ways. Hodent (2018) describes player motivation as part of self-determination theory (see Ryan & Deci 2000), meaning players are most motivated when their psychological needs are being met. Designers can shape the environment the player interacts with, meaning the actions and rewards they receive from them. For skill-based systems, like we are designing, each level is both a motivation and a reward. Acquisition of skills is meant to be timed across the gameplay, allowing steady rewards and a hook for players returning to the game (Hodent 2018, 137).

According to Hodent (2018, 65), players can also be motivated internally, by making sure they have a clear and manageable goal. Intrinsic motivation is at its best when the player is in a state of flow, meaning they are naturally motivated to play the game without the help of external rewards. An extrinsic reward can also be added, however, this can lead to destroying player's motivation, if the extrinsic reward is removed in later phases. This is due to confusion about the goal of the action, where an additional reward has replaced the internal reward (ibid. 66).

4.3 Balancing Tools

Many tools to develop balance are available with a variety of workflows and all of them are valid ways to balance a system or project. This section shares some common tools when balancing systems but will not go into detail or guide the use due to the scope of the thesis.

Spreadsheets

Spreadsheets are widely known as a go-to tool when dealing with statistics or numbers in general. Utilizing these tools allows designers to create formulas that depict progression in a skill tree and visualize them as graphs. It is common to use this tool to create numerical balance, which is often integrated into an engine using a script. Spreadsheets allow the creating of charts to visualize the balance statistically. Spreadsheets can for example be used in creating enemy loot drops. (Schreiber & Romero 2022, 696-698.) The drop

tables, like in *Break the Beyond* (Figure6, commonly contain an item, details related to it and its probability of being dropped.

	A		B	C		D
1	**Item Name**		**Item Level Index**	**Element**		**Weight**
2	Warhammer	▼	1	None	▼	15
3	Fire Element	▼	0	Fire	▼	5
4	Ice Element	▼	0	Ice	▼	5
5	Runic Ice Hammer	▼	0	Ice	▼	5
6	Paladin's Sword	▼	0	None	▼	5
7	Flail	▼	1	None	▼	5
8	Bardiche	▼	0	None	▼	5
9	Double Glaive	▼	0	None	▼	10
10	Bar Mace	▼	0	None	▼	2
11	Bow	▼	0	None	▼	8
12	Healing Potion	▼	1	Health	▼	2

Figure 6. Loot drop tables in *Break the Beyond*

Simulation tool

Simulation tools such as Machinations or Loopy are balancing tools that are useful in designing a game using data-driven design. By creating a systematic template, the economy of the game can be tested, while receiving preliminary data to drive further decision-making. The tool offers a more visual solution to balancing things like randomness in the game. All simulations can also be exported from *Machinations* to spreadsheets for a more traditional workflow (Machinations 2023). In the Break the Beyond game *Machinations* was used to clarify and balance the initial room-based level system, where the player moves from one room to another in random order and progresses in the game. This however was not an effective method for this development case, due to the payment plan and stage of the project. As the project's mechanics were bound to change in future, balancing bigger systems was not yet as relevant as in current public testing phase. While machinations could have been integrated to the game, it was not worth implementing due to lack of development resources.

Overall, the simulations allow feedback on the project's balance if its workflow has been integrated into a system. *Machinations* also has a plugin which allows live changes in the game loop and economy. The plugin allows

synchronization between *Unity* and *Machinations*. For example, designer can create the economy in machinations using different types of loops, create machine-based data, and edit variables live into the editor through synchronization (Figure 7).
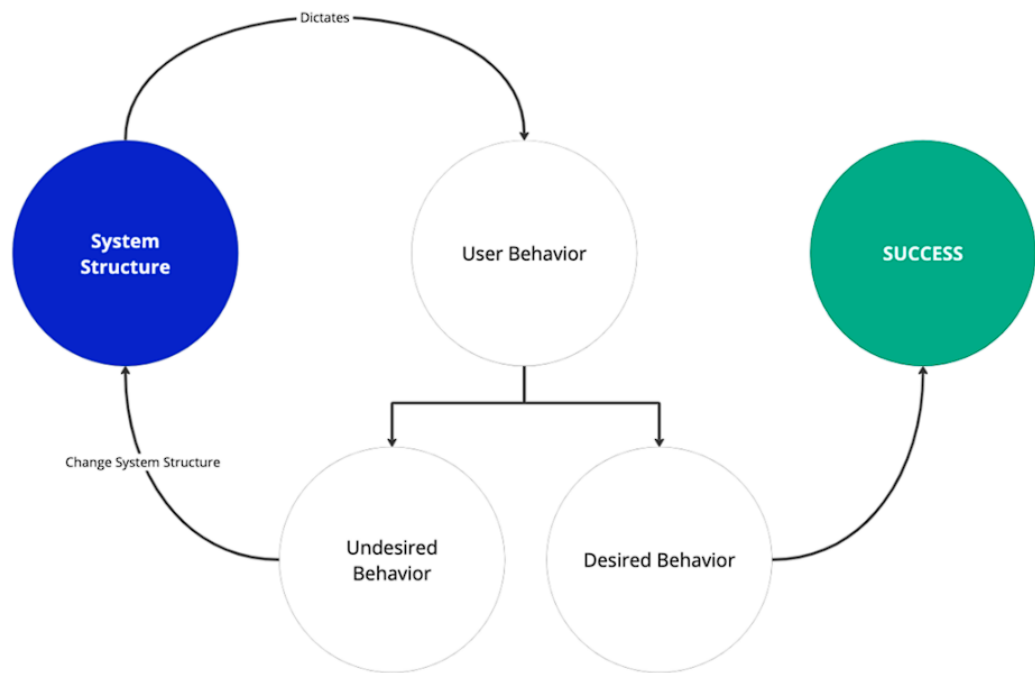

Figure 7. The structure of a system creates the behaviour of the system. (Machinations 2023)

Game Analytics

Game analytics allows the collection of quantitative data. There are several tools available but for this project one of the industry standards GameAnalytics was chosen. It is a tool for game developers, which collects the data through events placed in the game and arranges it in a visual and numeric data format, which is easier to interpret. Data can then be utilised by making more data-driven decisions for game development. The most common KPIs that are tracked using GameAnalytics for example are retention/churn, daily user acquisition or players' playtime. Just like machinations, it is used in data-driven design and can export data as spreadsheets.

5 SKILL TREE

Skill tree, also known as tech tree, is often visualized as a tree, which branches out to different passive or active abilities. The player can unlock their abilities during gameplay in exchange for their time, points, or other type of value. Skill trees are often used to engage and motivate players internally to play the game. Progression can be depicted in multiple ways. It can be story-based progression, where the player unlocks the story the more they play, or it can be an increase in the difficulty curve, meaning the player must be allowed an increase in power over time. This can be implemented with passive power-ups that are gained, better gear, or for example a skill tree.

In a skill tree, improvements are presented in the form of active or passive skills. Active skills are abilities that allow new and different gameplay styles for the player. These abilities give variety to players' gameplay, and in addition, motivate them to continue playing the game. Passive skills are usually upgrades that run in the background and enhance already possessed attributes and abilities, such as extra punch damage, or a bigger chance to dodge an attack (Schreiber & Romero 2021, 199). The skills can usually be unlocked in an order as players invest more time in the game. Skill trees motivate players intrinsically at a fixed rate. If the skills are communicated clearly and are meaningful, the player can use cognitive resources to perceive it as something they wish to attain. (Hodent 2018, 63.)

In the previous chapter about player experience (Chapter 3), it was discussed that designing for the player requires the mindset to be in player's position. Skills must also be meaningful to the player, as pushing them to make unmeaningful decisions will cause frustration and lessen the possibility of rewarding feeling. If a skill about inventory space is available in the skill tree, it should also be meaningful towards the end of the game, meaning the game must-have items that fill the inventory space. Otherwise, the skill might be overshadowed by a more epic skill. (Hodent 2018, 139.)

Skill trees can also be used to balance the pacing of the game in multiple dials mentioned in Chapter 4. Skill trees can be used to pace and time mechanics that are introduced to the player in one form or another. By restricting the

player's abilities from the start, the player is not overwhelmed by a variety of different abilities and combinations. Skill trees are also used to pace the overall progression of the game. For example, the game might be very hard to complete without a specific skill, which can be unlocked after playing a game for quite some time. This not only engages the player more but also motivates them to return and continue the game while exploring new ways to play the game (Game makers toolkit 2018). Skill trees can also be used as part of the player progression curve and are meant to pace and motivate the player with timely rewards and content. The most rewarding skills are abilities, that the player hopes to unlock. They give variety to gameplay and has an important role as being the goal player has set their eyes on.

6 USER TESTS

Testing is usually implemented as part of the development process in all its stages. It is implemented in almost every development team's workflow, which allows the best product possible. Playtests are referred to as physical user tests, where a member of the development team or third party performs a planned test on a player or test participant. Physical playtests, however, are performed before or during the beta phase of game development. The game is tested continuously for bugs, gameplay, or user experience problems by members of the development team and playtest participants.

The test contents are carefully planned in a test plan, which outlines many of the specifics of the tests, such as testing tools, how the test is performed, when and based on what. The test plan can also contain specific tasks called test cases, which are commonly a list of tasks to test. Agile development teams utilize shorter test charters, which are not as formal and contain information on what is the testing goal and how to reach it. Playtests are commonly performed in a small test group consisting of five to ten participants (Finska 2023, 11-12).

6.1 Conducting playtests

To iterate efficiently testing must be done to find problems in the design and to improve it. The tests are performed as an interview in two different groups.

One being at Xamk South-Eastern Finland University of Applied Sciences and another at LAB University of Applied Sciences.

The goal of the playtests is to figure out which version of the skill tree performs better in the pacing of progression and feeling of player motivation. While the interview is a more relaxed situation, it still drives players to play longer than in a real-life situation. Thus, players were asked to inform the test participant time when a player would normally stop playing. Players were also asked questions regarding the first two skills they would attain, such as if the amount of experience would feel reasonable. In the original test plan, it was also set as a goal to measure players feeling of reward, when gaining experience, however, due to development reasons, this was not included in the final test. At the time of performing the tests, the game does not have a visual indication of gaining experience from events, which is assumed to impact the feeling of reward.

During the interview, participants were informed about the stages of the test, which consisted of an introduction, tutorial, gameplay, questions, and conclusion. During the introduction, participants were informed shortly of the test and the purposes of the research. After the introduction, participants were given the game and played it for the duration of the tutorial. After the tutorial, the game was restarted, during which participants were asked questions. The game was restarted due to an immortality bug in the test builds, which was fixed by the restart.

During the tutorial and gameplay phase, the players were observed. Some of the players were also recorded to address them later in the interview. When a player received their levels, the author informed the test participants of attaining the level, and using the first level, as the intended skill-tree-related tutorial did not work in the test builds. After 30 minutes of gameplay, the sessions ended and the participants' experience points were recorded and questions in Table 4 were asked.

Table 4. Interview Questions

Question 1	How did you understand the skill system?
Question 2	How did you feel about gaining experience?
Question 3	How did you feel about gaining skills?
Question 4	How did you feel about the timing of the first and/or second skill?

Participants were thanked and invited to the following gameplay diary test mentioned below. After the tests were performed one of the participants in groups A and B was rewarded with a squeezable stress toy from Latchback Games. In addition to the physical AB test, a passive closed test was performed. For the test 12 participants were recruited and invited to the game's community discord server. The participants were given the current build of the game and were asked to play the game each day for the duration of the test. The build of the game was also tracked globally at the same time as the test, to gain more hard data. The participants were asked to fill out a short gameplay diary consisting of six repetitive questions for seven days (Appendix 1).

Test Group A – XAMK

The test was performed at the South-Eastern Finland University of Applied Sciences at Medusa studio. In total from 14 participants, eight of the test performers were either students or faculty. Participants ranged from 19- to 30-year-olds. Participants were acquired using advertisements in Discord and popups in campuses, to acquire the greatest number of participants. The initial form for registration to the research contained 11 registrations to the site, in addition to this three were recruited through events. Out of the 11 registrations to Kouvola, only seven participated in the research.

Tests were booked using the Koalendar booking service, which contained testing times ranging from 9 am to 5 pm for each day of testing. Each participant was asked the questions mentioned above either when filling out the test registration or before starting the test. A consent and data privacy

form were also signed, after reading the research information and privacy statement.

The focus of the test was the amount of experience participants were able to attain within 30 minutes of gameplay and their experience and preference for first skills. The participants were given an early version 0.3a_pre which was a version with the starting experience value for the game. This means that the first skill could be attained at 250 experience points, and the second at 1000 experience points.

Test Group B – LAB

Tests were performed mostly at Starthub Lahti office space and were aimed towards students and faculty of Lahti University of Applied Sciences. The participants ranged from 20 to 30-year-old participants. The number of participants from Lahti was smaller than in Kouvola. From five registered participants only three reserved a time, which meant the number of participants was short from originally planned. To achieve a more acceptable participant amount, some additional participants were gathered from the Kouvola region. In the end, there were six test participants in test group B, from which three had originally registered for the test. The rest of the participants were attained through social events to fill in missing participants.

The participants in group B were given version 0.3b_pre, in which the second skill of the skill tree was set to be unlocked at 750 experience points instead of the original 1000 experience points. The skill was changed to see if a lower amount feels more motivating than the 1000 experience points, forming a better hook for players.

6.2 Gameplay Diary

Participants were invited to join the game community discord platform and were given a role as test participants. Test participants were requested to install the game currently in the Google Play Store. The build at the time of the test was v.0.4.3, which had similar values to the build v.0.4a_pre used with test group A. The main difference with the build was additional features and

bug fixes unrelated to the thesis. Out of the 12 invited participants, only five joined the test.

The focus of the test was to gain more qualitative data similar to interviews, but on a more long-term basis as mentioned above in Chapter 3.2. During the test, participants were requested to fill out a gameplay diary consisting of six questions regarding their progression and gameplay experience of the game shown in Appendix 1.

6.3 Game Analytics

The version tracked was the same version (v.0.4.3) of the game as in the gameplay diary as it could potentially give more extensive hard data that could in theory support the results better. Players' actions can be tracked effortlessly, without disrupting players' game flow, which is not always possible with testing other methods mentioned earlier.

During the passive test, certain specific data points were tracked based on progression: the amount of experience gained within a session, the number of skills attained by a player, specific skills unlocked by a player and playtime by session. These custom data points were placed in the project by programmers of the game. In addition to progression related events, regular events such as retention are tracked.

6.4 Results

In both physical playtests, the initial hypothesis was that 60% of players would receive two skills in 30 minutes of gameplay. The hypothesis was proven to be false in both tests. All test participants managed to unlock the first skill within 30 minutes. None of group A managed to receive the second skill in the skill tree. In group B, four out of five participants managed to unlock the first skill and only one was able to unlock the second skill (Figure 8).

Group A – EXP by test participant
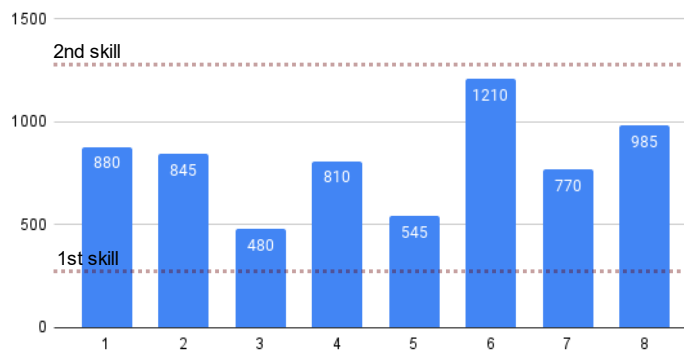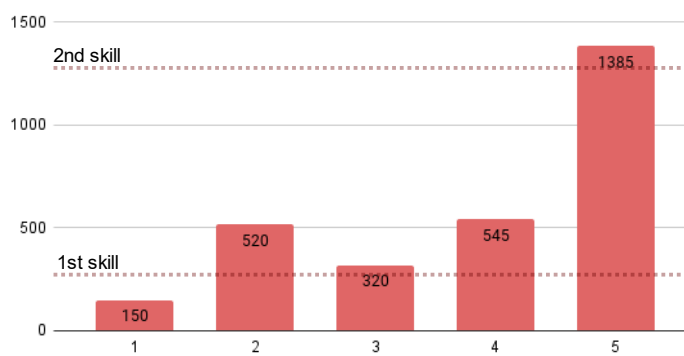

Group B – EXP vs Participant

Figure 8. Experience points by test participants in Group A vs. B

In regard to whether the timing of skills was right, participants in playtests A and B did feel like the first skill was well timed. Players earned enough experience to gain the first skill soon after the tutorial, in most cases after 10 minutes of gameplay. The second skill, however, was far away according to participants of group A, which also was indicated by zero number of participants unlocking it within the 30 minutes of gameplay. Test group B felt that the second skill was well timed overall, however was still difficult for a new player (Figure 9).

Is first skill timed well in Break the Beyond?
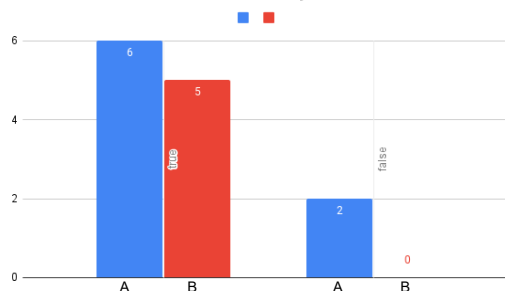
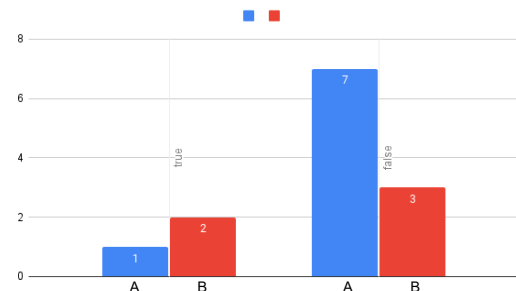Is second skill timed well in Break the Beyond?

Figure 9. Test group opinion of first and second skill timing

It could be concluded that the second skill needs to be paced lower than 750 experience points to also hook less proficient players. During the tests, some participants also felt the second skill should be made into an unlockable ability instead of a passive ability. In building the second skill increased the probability of being able to dodge an enemy attack by five percent. Players felt like skill should give the player the ability to dodge with a probability instead.

In physical playtests many of the players did not feel rewarded by the skill tree. When asked, many felt they did not feel rewarded due to not having a visual indication of received experience. However, this issue was known, but due to development schedule of the game, it was not in the scope of this thesis. Overall, the test gave important data about required changes to the early skill tree of *Break the Beyond* and gave insight into how it is perceived and used by the player.

In the gameplay diary all participants reported and wrote of their experience for the seven days. However, one of the test participants reported their game not working, and thus could not give information regarding most days of the experiment. This could have been addressed; however, the issue was not reported before the last day of the test. This is why player 5 is marked as an anomaly and is taken into consideration in calculations of mean averages. Originally, it was hypothesised that gaining a skill during the day would affect noticeably to the players' motivation to play the game for longer, however, according to the results players' overall experience did not change noticeably by the scale measured (Figure 10).

Figure 10. Player mood and skill gain per day

Based on the results it seems the amount of daily experience points gained increased after the first day, which could be an indication they progressed past the first realm: "Dungeon". This could be thanks to a gaining a skill, or a small boost in motivation to play the game longer (Table 5). More research might be needed to understand it better, as there is no way to compare the daily increase individually using the current hard data from the tests.

Table 5. Test participant's daily experience gained during the game diary test.

How much experience did you gain today?							
Day	Player 1	Player 2	Player 3	Player 4	Player 5	Daily Average	True Daily Average
1	310	540	290	145	0	257	199
2	380	1815	6000	710	290	1781	639
3	680	1050	2785	680	0	1040	482
4	1650	100	4000	285	0	1208	407
5	670	300	1315	1285	0	715	451
6	1355	1600	6265	565	0	1958	704
7	0	2000	2785	800	0	1118	560
Player Daily Average	721	1058	3349	639	41	1153	492

Compared to the average experience of 170 players of the same week using *GameAnalytics* between 16 and 22 October in Figure 11, the mean average of players experience was lower than the participants of the gameplay diary test (482 experience points). The passive tests average does seem to be close to some of the test participants daily average experience points by session if anomaly Player 3 and Player 5 is not considered. Player 3 was able to gain large amounts of experience points, indicating they played more sessions or longer sessions and Player 3 almost none due to issues mentioned above.
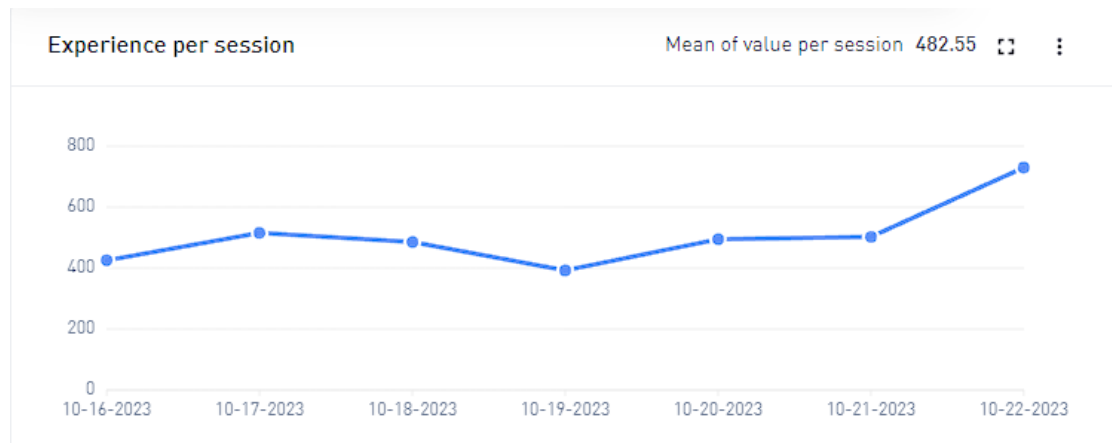

Figure 11. Experience per session 16 — 22 October 2023 (Latchback Games 2023)

However, if compared to a time before the specified timeline (10 — 17 October), when more players were active in the game and playing the same build, the average is very close to the results gathered from playtests and diary. According to the more extensive data presented in Figure 12, the 386 players recorded for the time were able to gather similar average of roughly 540 experience points per session.
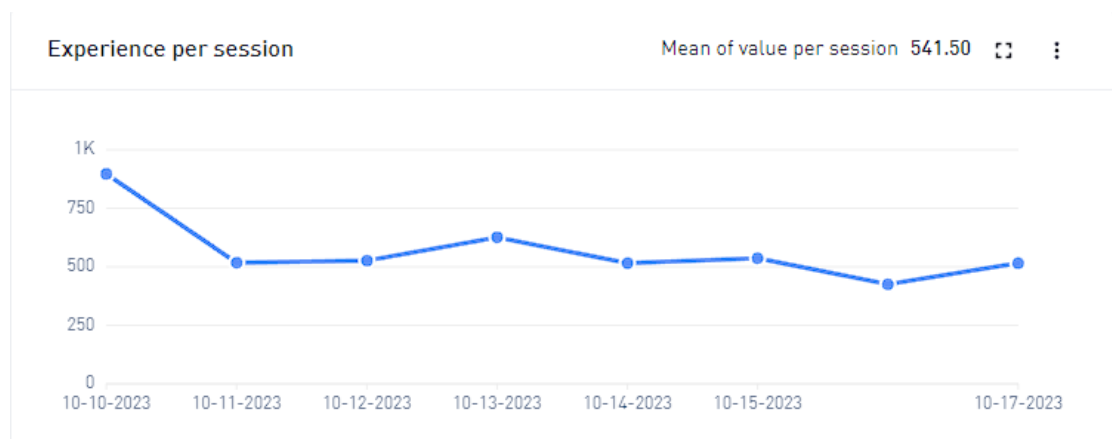

Figure 12. Experience per session 10 – 17 October 2023 (Latchback Games 2023)

While the game diary test did not specify the number of sessions during the day, it could be that players which surpassed the average amount of 540 played more than one session in a day. The physical playtests also support this as many of the players managed to reach the amount of experience points within a session as presented above. The difference between the test results in-game diary and analytics could also be affected by differences in retention and situation. According to gathered hard data, results are calculated from players, who have returned the game for three days, at which point most players have churned. Game diary test on the other hand was situationally framed to be seven days, meaning participants had a personal obligation to continue for the specified length (Figure 13).

Game Retention 10 – 17 October 2023

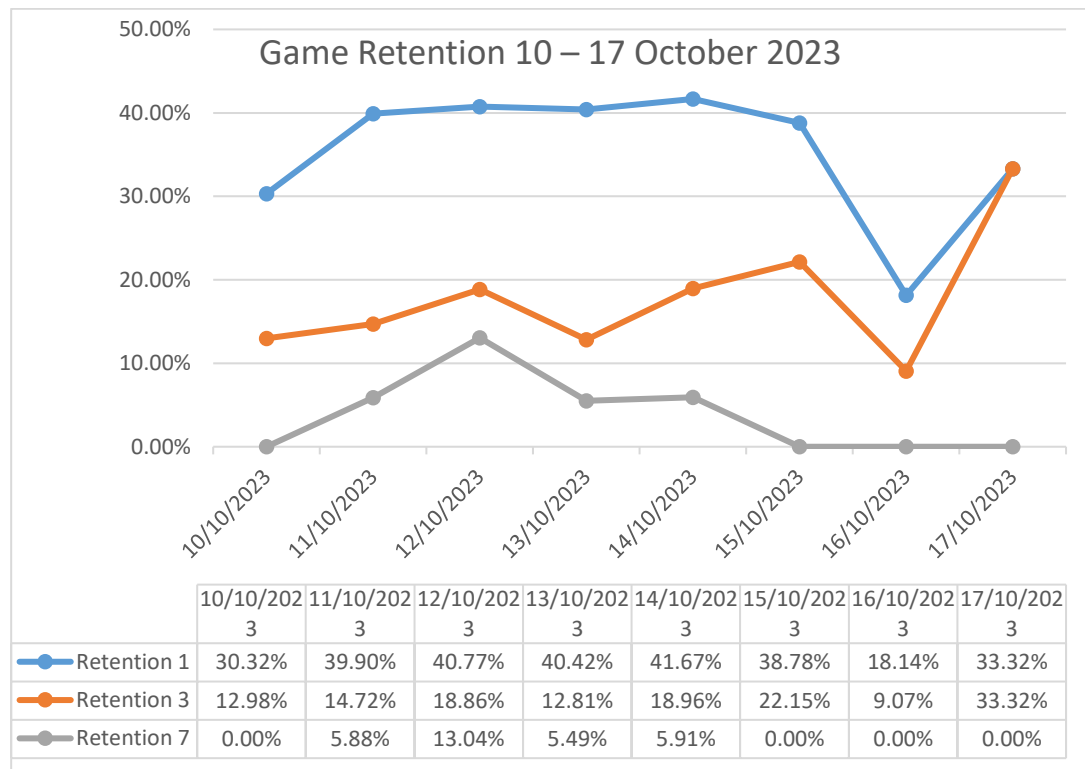|	10/10/2023	11/10/2023	12/10/2023	13/10/2023	14/10/2023	15/10/2023	16/10/2023	17/10/2023
Retention 1	30.32%	39.90%	40.77%	40.42%	41.67%	38.78%	18.14%	33.32%
Retention 3	12.98%	14.72%	18.86%	12.81%	18.96%	22.15%	9.07%	33.32%
Retention 7	0.00%	5.88%	13.04%	5.49%	5.91%	0.00%	0.00%	0.00%

Figure 13. Retention between 10 - 17 October 2023 (Latchback Games 2023)

During the tests, participants and players also reported many bugs related to both the skill tree and to game in general, which have been addressed in future builds after the thesis. The bugs will not be reported in this thesis as they are not relevant to the subject of the thesis.

7 CONCLUSION

This study was designed to answer the following questions: What should be considered balancing a skill tree in the *Break the Beyond* -game*?* And what pacing of players' progression is considered good? In the study, hard and soft data were used to clarify what and why is the player's experience and preferences in short- and long-term situations.

The main results showed that considering many dials of balance is important to create a consistent and well-paced experience. Players should feel both rewarded and engaged throughout the gameplay, with adjustments of difficulty and rewards in-order to be able to enter a state of flow. It is specifically important to consider the player's experience and the designer's intention, as knowing what drives the player's motivation can be crucial for a pleasant experience.

Furthermore, the research and research methods indicate that the skill tree in *Break the Beyond* requires adjustments, especially in the early part of the game to hook the players. The study however did not show results for skill-tree in the long term, as unlocking and balancing all the abilities is not yet possible due to low long-term retention.

The tests also pointed out multiple cases of improvement from bugs to changes in skills of skill tree. These changes have been reported to the development team during the writing of the thesis and are addressed appropriately in future builds of the game. The results support the development of the game and give directional data regarding next steps to balance the skill tree of *Break the Beyond.* As the development of the game has been continuous during the writing process, some of the results have already been used to improve the game.

In general, the findings suggest the skill tree should be easier to unlock in early states of gameplay, and skills should be more new gameplay mechanics, rather than supporting passive abilities. Some of the skills, but not all, may be passive upgrades that are meaningful to the player. The meaningful abilities however may change in the future. The hard data was

gathered from same build as game diary tests and show that the average experience points gained per session is ~500 exp. The results may vary due to the test situations. Physical playtests and Game Diaries gave soft data from a staged situation with a definitive time from one session to multiple sessions in a week. The hard data contains accurate statistics from a real-life situation, which means sessions can be affected by outside factors and players' motivation to play the game. The hard data suggests most players do not currently continue playing the game after three days due to a lack of motivation. The current soft data suggests a probable reason for this is lack of progression in skill tree, which is due to not gaining enough experience points in a session.

The results within this thesis are directional and do not reach statistical significance, which is why more comprehensive research is required in the future for more accurate results.

In conclusion, when balancing a skill tree designers should consider the pacing, difficulty, and reward of the game. Designers should seek to understand what and why players' experience is as it is, and to design the experience better. There are many tools to help with understanding the experience, but in the end, it all comes down to being able to understand players' motivation. Modern testing methods can be used to help understand the motivation, but the sample size should always be considered before making any definitive conclusions. The sample size of the performed tests was lower than initially planned. One way to prevent this could have been longer participant acquisition time, but it all comes down to authors available resources.

REFERENCES

Csikszentmihalyi, M. 2014. Flow and the foundations of positive psychology. London: Springer. E-book. Available at: https://link.springer.com/book/10.1007/978-94-017-9088-8 [Accessed 10 September 2023]

Game Makers Toolkit. 2019. Roguelikes, persistency, and progression. Youtube. Video clip. Available at:https://www.youtube.com/watch?v=G9FB5R4wVno&ab_channel=GameMaker%27sToolkit [Accessed 5 October 2023]

Game Makers Toolkit. 2018. Building better skill trees. Youtube. Video clip. Available at: https://www.youtube.com/watch?v=wsmEuHa1eL8&ab_channel=GameMaker%27sToolkit [Accessed 13 October 2023]

Finska, N. 2023. Modern Game Testing. Birmingham: Packt Publishing, E-book. Available at: https://ebookcentral.proquest.com/lib/xamk-ebooks/detail.action?docID=30607328&query=game+testing+modern [Accessed 5 August 2023]

Gazaway, D. 2022. Introduction to systems design. Addison-Wesley: Boston

Hodent, C. 2021. The Psychology of Video Games. Oxon: Routledge

Hodent, C. 2018. The Gamers Brain. Boca Raton: CRC Press

Jaffe, A. 2013. Understanding game balance with quantitative methods. University of Washington. Computer science and engineering. Doctoral dissertation. PDF document. Available at: https://digital.lib.washington.edu/researchworks/handle/1773/22797 [Accessed 4 September 2023]

Koster, R. 2013. Theory of fun for game design. 2nd ed. O'Reilly Media: Sebastopol

Kramarzewski, A. & De Nucci, E. 2018. Practical Game Design. Birmingham: Packt Publishing. E-Book. Available at: https://ebookcentral.proquest.com/lib/xamk-ebooks/detail.action?docID=5353673 [Accessed 19 October 2023]

Lankoski, P. & Björk, S. (eds.) 2015. Game research methods. ETC Press: Pittsburgh. E-book. Available at: https://press.etc.cmu.edu/books/game-research-methods

Machinations. 2023. Machinations S.àr.l. Web page. Available at: https://machinations.io/ [Accessed 13 November 2023]

Ryan, R. & Deci, E. 2000. Intrinsic and Extrinsic Motivations: Classic Definitions and New Directions. Murphy, P. (ed). Contemporary Educational Psychology. New York: Elsevier, 54–67, Available at: https://www.sciencedirect.com/science/article/pii/S0361476X99910202?via%3Dihub [Accessed: 13 November 2023]

Schreiber, I. & Romero, B. 2022. Game Balance. Abington: CRC Press. E-book. Available at:https://ebookcentral.proquest.com/lib/xamk-ebooks/detail.action?docID=6648067 [Accessed 27 July 2023]

Schell, J. 2020. The art of game design. 3rd ed. Boca Raton: CRC Press

Hirsjärvi, S., Remes, P. & Sajavaara, P. 2009. Tutki ja kirjoita. 15th ed. Tammi: Jyväskylä

Zubek, R. 2020. Elements of Game Design. MIT Press: Cambridge. E-Book. Available at: https://ebookcentral.proquest.com/lib/xamk-ebooks/detail.action?docID=6270462 [Accessed: 1 November 2023]

LIST OF FIGURES AND TABLES

Figure 1. The core loop of Break the Beyond. Suojanen, J. 10 Nov 2023.

Figure 2. Break the beyond. Latchback Games. 2023. Available at: www.play.google.com/store/apps/details?id=com.Latchback.BtB&hl=en&gl=US [Accessed 17 Nov 2023]

Figure 3. Framework of thesis. Suojanen, J. 5 Oct 2023.

Figure 4. Thesis timetable. Suojanen, J. 5 Oct 2023.

Figure 5. Different types of intermittent rewards and their approximate impact on behaviour. Hodent, C. 2018. The Gamers brain. E-book, Boca Raton: CRC Press. Available at: https://ebookcentral.proquest.com/lib/xamk-ebooks/detail.action?docID=4943930 [Accessed: 17 Nov 2023]

Figure 6. Loot drop tables in break the beyond. Suojanen, J. 15 Oct 2023.

Figure 7. The structure of a system creates the behaviour of the system. Alexandre. 25 Oct 2023. Machinations

Figure 8. Experience points by test participants in group A vs. B. Suojanen, J. 13 Nov 2023.

Figure 9. Test group opinion of first and skill timing. Suojanen, J. 13 Nov 2023.

Figure 10. Player mood and skill gain per day. Suojanen, J. 13 Nov 2023.

Figure 11. Experience per session 10 – 17 October. Latchback Games. 13 Nov 2023.

Figure 12. Experience per session 16 – 22 October. Latchback Games. 13 Nov 2023.

Figure 13. Retention between 10 – 17 October 2023, Latchback Games. 13 Nov 2023.

Table 1. Research methods of thesis. Suojanen, J. 13 Nov 2023.

Table 2. Research methods pros and cons. Suojanen, J. 10 Oct 2023.

Table 3. Different rewarding intervals. Zubek, R. 2020. Available at:
https://ebookcentral.proquest.com/lib/xamk-
ebooks/detail.action?docID=6270462 [Accessed: 1 November 2023]

Table 4. Interview Questions. Suojanen, J. 2 Oct 2023.

Table 5. Test participant's daily experience gained during the game diary test.
Suojanen, J. 2 Oct 2023

Monday

How was your gameplay today?

Please let us know how your overall mood is today after playing the game.

	1	2	3	4	5	
Very frustrating	○	○	○	○	○	I loved it!

Was there a particular reason for the previous answer?

Your answer

How much experience did you gain today?

You can track your experience amount by opening your skill tree.

Your answer

How many skills did you unlock today?

- ☐ 1
- ☐ 2
- ☐ 3
- ☐ More than 3

Are you aiming to get a particular skill? Why?

Your answer

Did any skill feel useless? Which and why?

You can see skills by opening the skilltree and tapping the skill in game.

Your answer