

Irina Kekman (2109430, GDKV21SP)

RANDOMIZING LINEAR LEVELS

Exploring the Impact of Randomization
on Player Experience

Bachelor's Thesis

Bachelor of Culture and Art

Degree Programme in Game Design

2025



South-Eastern Finland
University of Applied Sciences

§

Author (authors)	Irina Kekman
Thesis title	Randomizing Linear Levels: Exploring the Impact of Randomization on Player Experience
Commissioned by	MiTale Oy
Time	2025
Pages	39 pages, 3 pages of appendices
Supervisor	Suvi Pylvänen

ABSTRACT

The research conducted and described as a part of this thesis investigates how randomization affects player experience in traditionally linear level design. By introducing randomized elements into the dungeon environment, the study aims to understand how alterations in room order and visual variety influence players' experiences, particularly, their motivations for playing, inclination to explore, and overall engagement.

The Willow Guard project and research center around a custom-built game prototype with and without applied randomization. Two versions of a dungeon level were created: one strictly linear and the other featuring various randomized elements preserving narrative pacing and gameplay goals. The study aims to assess how these structural changes impact player perception, immersion, and navigation.

The findings suggest that even subtle randomization in linear levels can enhance perceived replayability and curiosity, though it may also introduce disorientation or reduce narrative cohesion if not carefully balanced. These insights offer valuable guidance for game developers seeking to integrate dynamic systems into structured gameplay without sacrificing clarity or resources.

Keywords: level design, linearity, procedural generation, randomization

CONTENTS

1	INTRODUCTION	4
2	APPROACHES TO VARIABILITY IN LEVEL DESIGN.....	6
2.1	Linearity	6
2.2	Procedural Generative Algorithms	7
2.3	Environmental Design.....	9
2.4	Previous Research	10
3	WILLOW GUARD PROJECT	12
4	RESEARCH DESIGN	15
4.1	Initial Research	16
4.2	Hypothesis	17
4.3	Qualitative A/B Testing	18
5	TESTING BUILD IMPLEMENTATION.....	19
5.1	Dungeon Generation	20
5.2	In-room Randomization	21
6	RESEARCH FINDINGS.....	25
6.1	Dungeon Layout	26
6.2	Collectible Items	27
6.3	Environmental Variations.....	27
6.4	Exploration.....	28
6.5	Key Findings	29
7	DISCUSSION AND ANALYSIS	31
7.1	Comparing Findings to Previous Research	31
7.2	Recommendations for Level Design in Future Development of Willow Guard.....	32
8	CONCLUSION.....	34
	REFERENCES	36

1 INTRODUCTION

The continuously increasing number of games released annually intensifies competition, forcing developers to differentiate through innovative features, enhanced gameplay mechanics, and unique storytelling elements. Developers strive to retain players through implementing game features that enhance replayability of the game, such as random or procedural content generation, story branching, or other dynamic elements. Large corporate entities, such as acclaimed AAA developers, are generally believed to possess sufficient resources to implement varying degrees of randomization and procedural generation, allowing them to create a highly unique experience for each player. However, smaller companies or those with limited budgets need to more carefully evaluate the efficiency of each approach balancing the time consumed for implementation against the impact on the players.

There are several factors that impact players' immersion, one being previous experience with games and expectations. This factor is completely out of game developers' control; however, game mechanics and game content remain key aspects that developers can influence to enhance immersion. Game mechanics define the rules and systems governing player interactions, while game content includes elements such as narrative, visuals, and audio, all of which contribute to player's sense of presence and other experiences in the game world.

The research for this thesis explores the impact of variability within levels on the players' experiences with and perception of variation in Willow Guard, a game developed by MiTale Oy. The findings of this study will be utilized by the commissioning company in decision making for the further development of Willow Guard and its potential additional downloadable content.

This research examines the challenge of designing vast amount of content for the game as well as diverse levels with a certain degree of variation, investigating how changes in level design affect players' perception of both in terms of the gameplay as well as aesthetically. The study aims to evaluate the extent of this impact and provide reliable data for future level design and data-driven

development decisions. Thus, the main research question stands as follows: Are players able to perceive increased variation of the levels' contents?

Specifically, this study assesses the impact of increased variability through randomization of key level elements, such as loot distributions, environmental elements, and dungeon layout, on player perception and immersion. It seeks to determine whether players recognize and value variation in visual elements and item placement within levels. Understanding this relationship will help assess whether the benefits of randomization justify the development resources required for its implementation. Thus, the secondary research question is: Does perceived variation in levels effect players' immersion in the game?

Additionally, this study explores the balance between the effort invested in randomization during level design and the perceived benefits for players. This includes examining whether players respond positively to increased unpredictability or if familiarity, through repeated exposure to static patterns, enhances engagement by fostering a sense of mastery. The objective is to determine whether extensive variability is necessary or if player preferences and genre conventions mitigate its importance.

One of the primary constraints of this research is time, which impacts its implementation in several ways. First, the recruitment period for playtesters is limited, as is the amount of time they are willing to dedicate to testing the game. Consequently, the duration of each playtest session and the number of playthroughs per tester must be carefully considered to ensure meaningful data collection while respecting participants' availability.

Another limitation that may influence the observations and results, is the sample size. Larger testing groups typically yield more statistically significant and reliable data, whereas a smaller sample size may introduce biases or limit the generalizability of the findings.

2 APPROACHES TO VARIABILITY IN LEVEL DESIGN

This chapter introduces the theoretical background and explains various approaches used in level design to diversify the contents of the game, increase replayability, and enhance players' experiences.

2.1 Linearity

A level in a video game may be compared to physical space in the real world. Like real-world locations, levels can be designed to contain smaller spaces, structured in various ways depending on the designer's objectives, however, relieved from the constraints of the physical world (Totten 2019). These structural choices influence how players understand and experience the environment, for example, an intersection would be more interesting than a straight road, and furthermore, intersection leading to various outcomes would even require decision making (Totten 2016). A common classification of level design approaches distinguishes between linear and non-linear level design, oftentimes including hybrid models, the latter often referred to as semi-linear level design (Salmond 2021).

In a linear level, player progression is predetermined by the designer, guiding players from the starting point to the objective along a singular path through the level (Salmond 2021). While this approach may include occasional branching paths, secret areas, optional challenges, or, in other words, an illusion of choice, the overall design ensures that players experience key events in a specific sequence (Wilson 2018). This structure offers benefits such as clear control over pacing and narrative progression. However, it can also feel restrictive, limiting player agency and reducing the sense of exploration, leading to players abandoning the game if they get stuck. (DeLally 2020.)

In contrast, non-linear level design provides players with multiple paths and choices, accommodating different playstyles and encouraging exploration, which can lead to varied gameplay experiences and game endings, as implemented in such classic examples as Fallout series and Arcanum: Of Steamworks and

Magick Obscura (Miszczyk 2017). The most extreme form of non-linear level design is found in open-world games, where the entire game world serves as a playable area, players' movement is not confined by the level boundaries rather guided through storytelling, environmental landmarks, or enemy placement (Salmond 2021).

Many modern games, however, integrate elements of both linear and non-linear design, creating so-called semi-linear level structures, as defined by Kremens (2009). In these systems, players are granted a degree of freedom to explore and encounter game events in a flexible order while still being directed toward key moments, known as bottlenecks. These bottlenecks serve critical functions, such as ensuring players experience essential story elements, acquire important items, or face major challenges such as boss encounters. (Kremers 2009.)

2.2 Procedural Generative Algorithms

Handcrafting levels for a game is a time-consuming task alternative to which is utilization of the procedural content generation algorithms. Procedural content generation (hereinafter referred to as PCG) is the creation of the game contents by a preprogrammed algorithm with limited or indirect input from the designer yielding results that are perceived by players to be random. (Sampaio et al, 2017.)

True randomness in mathematics and computer science, however, is a rare occurrence. The well-known mathematician of the 20th century John von Neuman (1951, 36-38) has written:

“Any one who considers arithmetical methods of producing random digits is, of course, in a state of sin. For, as has been pointed out several times, there is no such thing as a random number, there are only methods to produce random numbers, and a strict arithmetic procedure of course is not such a method.”

True random number generators (abbreviated as TRNGs) are inherently hardware devices that utilize physical processes, and the true random numbers

produced by these TRNGs are used in various simulations, industrial testings, cryptography, and other demanding spheres (Stipcevic & Koc 2014).

In contrast to TRNGs, pseudorandom number generators (or PRNGs) are deterministic algorithms that produce seemingly random number sequences and are in fact predictable. One way to exemplify pseudo-random number is to establish a formula that will define a long string of seemingly random numbers and then use a seed to pick or generate from this string. The “seed” is a number which is fed into the PRNG and which changes something in the algorithm, logically changing the result as well, while the same seed fed into the PRNG multiple times will always yield the same result. (DigiDigger 2020.)

Pseudorandom numbers are widely used in various PCG algorithms, being fundamental tools for introducing variability and unpredictability into the generation process. One such common algorithm is noise and its most popular variation Perlin noise (an example of visualized noise is presented in Figure 1 below), which is utilized in game development for generating, for example, terrains (Sainio 2023) and roads (Ginting et al. 2019).

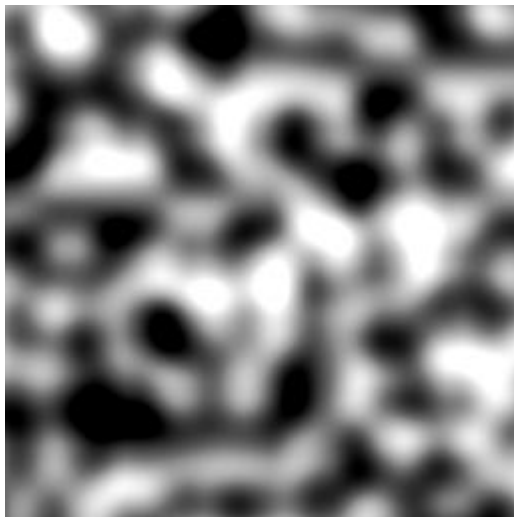


Figure 1. An example of Perlin noise rendered using TouchDesigner

Another common algorithm used for procedural generation is named cellular automata. The game space is divided into sectors or cells with their respective coordinates and each cell is defined by one of at least two states. A set of rules defines how the cells change their states and evolve each iteration. (Laitaruk and

Hryshanovych, 2024.) This algorithm is particularly noteworthy, as it has been implemented as a standalone game named Game of Life, which relies solely on the inherent rules and behavior of the cellular automaton. The game was developed by mathematician John Horton Conway, and is in essence the cellular automaton algorithm itself, whose evolution is determined by its initial state, follows the rules, and does not require additional input. (Roberts, 2015.)

There are multiple other methods and algorithms of procedural generation in addition to the abovementioned, for instance, described in the article by Laitaruk and Hryshanovych (2024): graph grammar, genetic algorithms, physics simulation, to name a few. These methods are complex and are applicable to more sophisticated game systems and worlds, oftentimes 3-dimensional worlds.

Random or randomized generation in game design is a notable case of game content generation, which does not, despite the naming, use TRNGs. The randomized generation term is used to describe populating game space with predesigned objects that are selected and pulled from a pool of objects and placed in the level based on a scripted set of rules (Bycer, 2018a). It is noteworthy particularly because it is the method used for adding variability within the Willow Guard levels for this research.

2.3 Environmental Design

Level design and environmental design terms are used synonymously in some sources (Bycer, 2018b) since visual aspects of the game world contribute to the game narrative and mechanics through contextually rich and aesthetically coherent levels (Porokh, 2023).

An important factor for visual variation in game levels is the presence of biomes, distinct environmental zones that often contribute to worldbuilding and narrative through visual differentiation and thematic cues. In some games, biomes are not purely aesthetic but introduce unique gameplay mechanics or traversal challenges. The shifting of biomes provides players with a natural sense of progress and novelty. A diverse set of environments can serve as a strong

motivational factor for players to explore large game worlds, especially when each biome introduces new visual and contextual elements to discover. (Geniuscrate 2024). Environmental storytelling further enhances the player experience by embedding narrative cues directly into the space, allowing players to interpret and piece together the story through visual details rather than explicit exposition. Rather than being told what happened, players uncover meaning through clues such as the placement of objects, environmental damage, or background details. This approach fosters immersion and personal interpretation, resulting in a deeper connection to the game world.

Lastly, room types within levels also contribute to environmental variability. The ability for players to quickly identify a room's function, be it a combat arena, puzzle area, or loot chamber, adds clarity and rhythm to exploration. When such functions are communicated visually through environmental cues, players develop a better understanding of spatial logic and level flow, which can enhance both engagement and navigational satisfaction (Geniuscrate 2024).

2.4 Previous Research

The findings of most research on game development are difficult to transfer from one subject (game) to another due to natural discrepancies in game structures, study technological constraints, methods used to gather data, or for other reasons, and should be taken into consideration with qualifications. Multiple articles described hereinafter (Conner et al. 2017; Rodriguez & Brancher 2017; Bastos et al 2017) have explored the relationship between level design and immersion, aiming to understand how different level design approaches, such as static versus procedurally generated levels, affect player experience.

Two studies stand out as particularly relevant to the exploration of level variation and immersion in game design. The first study by Conner et al. (2017) examines the differences in immersion between procedurally generated and human-crafted levels. Their research found that levels designed by human developers achieved a higher degree of player immersion compared to purely procedural environments. This suggests that while procedural generation can introduce

novelty and replayability, handcrafted levels tend to be more cohesive, structured, and emotionally engaging. (Connor et al. 2017.)

The second relevant study by Rodrigues and Brancher (2017) investigates the impact of random versus static level design in a mathematical game. Unlike the previous study, this research found no significant difference between procedurally generated and static levels in terms of player experience across multiple metrics. The only noticeable difference was that players in the static level version sought more explanations for game elements, potentially indicating a difference in how predictable or intuitive the levels felt. This finding suggests that the role of procedural generation in immersion may depend on the genre and complexity of the game. In a game with simple mechanics, such as a math-based game, level variation may not significantly influence immersion, whereas in more narrative-driven or exploration-based games, handcrafted levels might provide a more engaging experience. (Rodrigues & Brancher 2017.)

The impact of graphics on immersion, among other aspects, has been studied by Bastos et al. (2017) yielding the following results. Relevant and suitable to the game world and gameplay situation graphics help to achieve the feeling of presence through disbelief suspension. As authors refer to them, “credible” graphics cohesively fit into the gameplay mechanics and story. Although this research does not cover non-linearity or procedural content generation, it is nevertheless relevant to this research on Willow Guard level design, as the main element that is randomized, and which visibility and impact are evaluated.

The contrast between these two studies highlights the need for further investigation into how different types of procedural generation influence player perception. This research examines a middle ground between purely handcrafted and fully procedural levels, introducing a degree of randomness while maintaining a structured design. By comparing player immersion between levels with varying degrees of procedural generation, this study aims to determine whether carefully controlled randomness can enhance engagement or if a more handcrafted approach remains superior. The findings will contribute to ongoing discussions

about the balance between dynamic content generation and deliberate level design in optimizing player immersion.

The selected studies collectively underscore the complexity of defining immersion in relation to level design, revealing that its effectiveness is highly dependent on game context. Conner et al. (2017) demonstrate that handcrafted levels tend to promote deeper immersion, likely due to their narrative coherence and intentional design, while Rodrigues and Brancher (2017) suggest that in mechanically simple games, such as mathematical puzzles, level variation has limited impact on player experience. Bastos et al. (2017) further shift the focus by highlighting the importance of graphical credibility in fostering a sense of presence, which, although not tied directly to procedural generation, reinforces the idea that visual consistency enhances immersion. Together, these studies indicate that there is no universal model for achieving immersion through level design; rather, its success is shaped by the interaction between visual fidelity, gameplay complexity, and structural predictability. This variability signals a need to approach each game on its own terms when investigating how design choices affect engagement.

While the cited studies offer valuable insights into the relationship between level design, immersion, and graphical coherence, their findings cannot be directly applied to Willow Guard due to fundamental differences in game nature, genre, and testing conditions. The games studied differ significantly in mechanics, narrative depth, and player interaction models, all of which shape immersion in context-specific ways. Therefore, a tailored investigation is necessary to explore how controlled randomization within Willow Guard's unique hybrid structure influences player experience. This research acknowledges prior work but positions itself as a complementary study focused on the nuances of its own design environment.

3 WILLOW GUARD PROJECT

Willow Guard blends elements of action RPGs (Figure 2) and deck-building (Figure 3) games with a branching interactive narrative (Figure 3). By the time of

this research, the branching narrative has been defined to be one of the core features of the game. However, applying other approaches to implement variability remained a matter of efficient resource allocation. To support data-driven decision-making, the developers chose to explore how varying degrees of randomization for selected in-level elements impact players' experience with the game.



Figure 2. Action aspect of Willow Guard



Figure 3. Deck-building aspect of Willow Guard

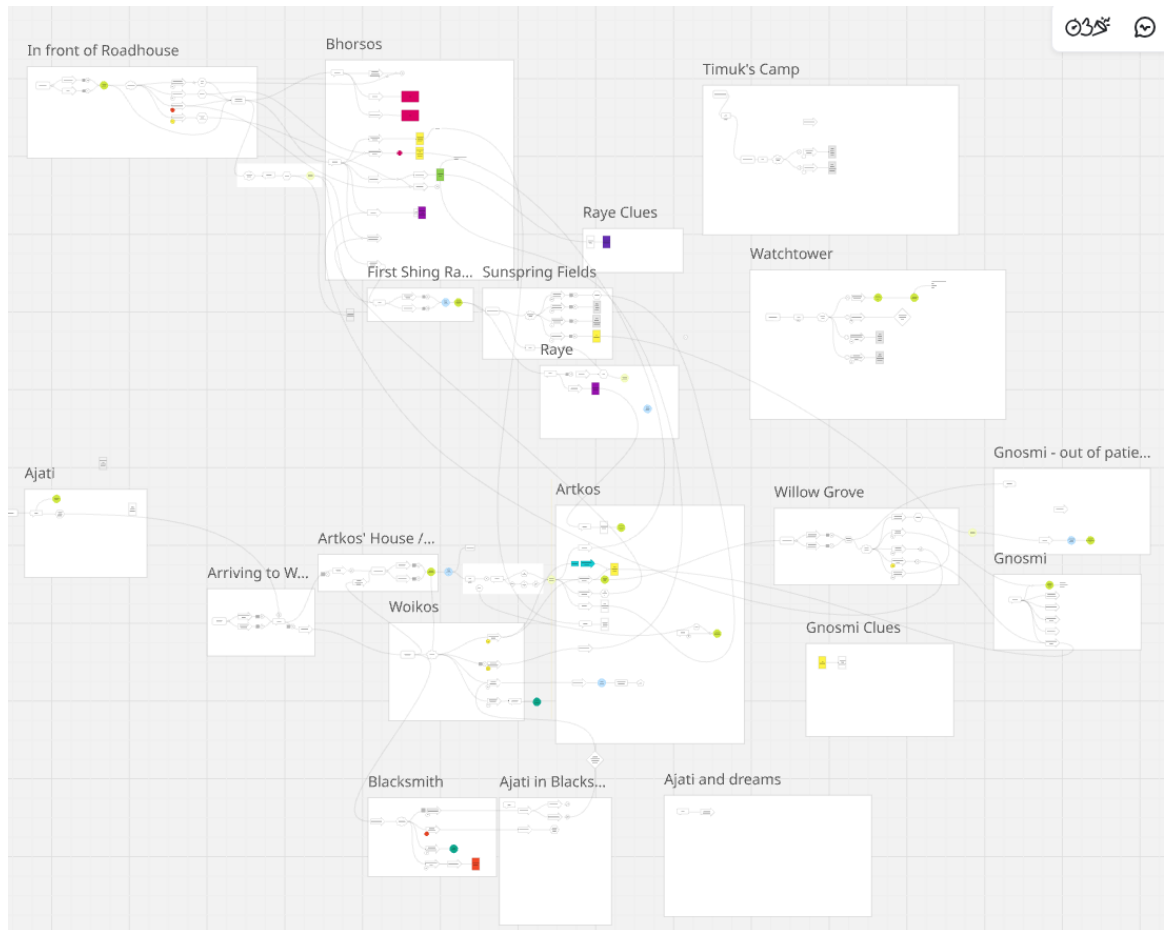


Figure 4. Willow Guard dialogue branching scheme (early-stage development)

To explore the effects of additional randomization of dungeons, it was decided to develop a focused A/B testing process. This test aimed to compare two different versions of one game level, allowing for a structured investigation into how players perceive and respond to varying degrees of environmental and item-based variability. The following section describes the steps made to carry out this test and collect relevant data to inform further development.

For the purposes of this research an overview of existing levels, their contents, current state of variation, and inherent visual aspects were evaluated, and one specific level, Fields, was selected for testing. This choice was deliberate rather than random, based on two key criteria. First, the Fields level is revisited multiple times throughout the game, which raises the likelihood of players experiencing visual fatigue over time. Second, as a naturally uniform, field-based environment,

it offers limited inherent variety in its environmental elements. These characteristics make Fields a particularly suitable candidate for evaluating the effectiveness and perceived value of added variability through randomization.

To introduce variability within the existing rooms and assess its impact on player perception, a new script was implemented to randomize certain elements within the level. The detailed description of the test builds and elements randomized within follows in Chapter 5. Shortly, in randomized version of Willow Guard, procedural generation is used selectively. Rather than fully relying on complex procedural systems, the game applies modular randomization to specific elements, particularly dungeons, loot placement, and environmental details, while maintaining handcrafted aesthetics and narrative flow.

4 RESEARCH DESIGN

This chapter outlines how the research used to evaluate how randomization in linear levels affects player experience was designed. The research methods, test structure, and data collection approaches are detailed in the following sections. Figure 5 below reflects the main stages of the research and their implementation time.



Figure 5. Research project timeline

Prior to commencing the testing itself, a comprehensive overview of the current state of the existing levels was conducted. It included analyzing their existing structure, visual composition, visual distinctiveness of various areas, and potential improvement from randomization.

Based on the internal discussions among developers, findings from the initial research, and the overall design goals of the project, hypotheses were formulated. These hypotheses aim to address specific questions, such as whether randomization introduced into levels would increase perception of

variation in players, and whether players would consider this variation to have a positive impact on their experience.

Two distinct versions of the level are created: a control version (A), which maintains a static layout, and a variant (B), which introduces randomization in room sequencing, loot distribution, and selected environmental elements. These variations are designed to isolate the effect of procedural changes while keeping other factors constant

4.1 Initial Research

The initial phase of the research focused on identifying the core needs of the indie development team, particularly regarding the degree of variability necessary to maintain player engagement. It included discussions with the developers on the elements to be randomized and the number of levels or combinations of randomized elements required for the players to perceive the variation.

This stage also involved formulating the main research questions, whether variation in linear level design is noticeable to players, and if so, whether it effects their immersion, as well as identifying secondary considerations, such as the potential impact of variation on replayability and visual fatigue. Discussions with the developers highlighted a practical question about how to introduce meaningful variation without drastically increasing production scope.

The evaluation centered on several types of in-level variation: gameplay-relevant elements like placement of loot, heal items, and resources, environmental features that shape the level layout, and purely visual changes such as terrain textures. These categories were considered both independently and in combination. Additional attention in the initial research was paid to the fact whether certain rooms or areas appeared too distinct, potentially becoming overly memorable and undermining the intended sense of variation, or otherwise, whether certain biomes were inherently repetitive potentially benefiting from introduction of variability.

4.2 Hypothesis

The development team formulated two primary hypotheses regarding the potential impact of randomization of environment and collectible items on immersion. These hypotheses reflect the contrasting playstyles observed among different target audience demographics, ranging from fast-paced, goal-oriented players to those who favor exploration and narrative depth.

The first hypothesis suggests that players, particularly hardcore gamers and speedrunners, are primarily focused on combat and gameplay mechanics, paying little to no attention to the surrounding environment during intense encounters. These players prioritize efficiency, often optimizing their strategies to complete levels as quickly as possible. As a result, environmental details may be perceived as inconsequential, making any changes to level layouts or visual elements negligible in their overall experience. If this hypothesis holds true, then efforts to introduce environmental randomization would have minimal impact on player immersion for this subset of the audience.

The second hypothesis states that players, particularly those invested in narrative-driven experiences, do explore and engage with their surroundings, paying attention to environmental storytelling and visual variety. For these players, dynamic and varied level design may enhance immersion by reinforcing the sense of discovery and unpredictability within the game world. If randomization creates a perception of a more organic, living environment, it could contribute to deeper engagement and reduce visual fatigue, particularly in levels that are revisited multiple times.

This study seeks to determine which of these hypotheses holds true by examining player behavior and feedback in response to randomized and static level designs. The results will provide valuable insights into whether level variability meaningfully and positively impacts immersion and whether different player demographics exhibit distinct preferences regarding level design approaches.

4.3 Qualitative A/B Testing

An A/B test, also referred to as split test, or randomized controlled trial, was decided to perform on two groups of testers with subsequent feedback in the form of questionnaires and personal interviews. Split test is a useful tool for comparing two or more versions of a product in order to receive data for further informed decision-making. This method is widely used in the game industry to test a variety of game aspects, from placement of UI elements to microtransaction prices. Comparing responses to version A of a game build to version B, developers are able to determine which model is more effective (Kekki, 2024.)

Two different game builds were distributed to two groups of playtesters. The game builds were uploaded to Steam platform as branches accessible with passwords, different for each build. Testers were given passwords and instructions on how to access their respective builds. The first group was assigned to test the static dungeon (Test Build A), while the second group was assigned to test the dungeon with an updated randomization system, which included variations in environment and loot placements, as well as the starting dungeon layout (Test Build B). Testers were given an average of a week time to test the game and reply to the feedback questionnaire. Altogether 18 testers from Finland have been involved in the testing, nine for each test build. Details on the test builds implementation and their differences are presented in Chapter 5.

The feedback questions were designed and formulated for the respondents to not be guided into a certain way of analysing the playtest experience, rather to reflect on their personal preferences and approaches to completing the level. The questions did not point out whether the tester has played a static or randomized version, or even the nature and the purpose of the tests in general but facilitated analyzing what guided them through the level and elements of the game impacted the decision making.

The questionnaires for both groups were identical with the exception for one question regarding the starting layout. The layout was pointed out and questioned

about because it was the only one potentially identifiable randomized element, and asking about it specifically would not have caused respondents to overanalyze the testing goals or build structure and thus would not hinder their answers to the further questions. The questionnaires can be found in Appendix 1.

Traditional A/B testing is a quantitative research method collecting objective quantifiable data. For the purpose of this thesis, heatmaps were considered as one of the possible ways to collect objective data from testing; however, it was defined that the battle mechanics would skew the results. Based on internal gameplay testing, it was defined that due to the time of level playthrough being approximately 80% battle and 20% exploration, exploration data would be insignificant and unidentifiable from the battle data. The data collected through the current research is naturally subjective, making this A/B testing effectively qualitative.

5 TESTING BUILD IMPLEMENTATION

Prior to the research, a simple random generation algorithm of connecting premade rooms has been used in Willow Guard without any randomization of elements within the rooms. This caused an issue of identical rooms being generated next to each other. This issue has been decided to approach via introducing randomization of specific groups of items inside each room. This chapter describes in detail how variability was implemented through randomization of various level aspects, such as collectible items (loot, gold, or items increasing health or stamina), and environmental elements.

Two game builds limited to one area (fields) have been compiled for the players to test and assess their experience. In both builds, the dungeon contained three different basic rooms, three different aegis rune rooms, one boss room, and one puzzle room. The script in Test Build A always generated the same dungeon layout, with all collectible items placed in the same spots, and without any changes to the room contents. Test Build B offered players dungeons with randomized layouts, changing loot placements, and an array of randomized

environmental elements. Table 1 reflects the elements and their static or random nature in the respective builds.

Table 1. Test builds comparison

	Test Build A (static)	Test Build B (randomized)
Dungeon layout	static	random
Loot placement	static	random
Environmental elements	static	random
Enemeis spawn	random within static spawn area	random within static spawn area

The script managing enemies spawn location remained the same in both builds, spawning monsters at a random spot within their dedicated circular areas.

5.1 Dungeon Generation

The dungeon system in Willow Guard is based on a modular architecture of premade rooms, each designed with exits that connect in the cardinal directions: North, South, East, and West. When a dungeon is started or restarted, an algorithm selects and connects a sequence of these rooms to form a complete path, ensuring the required number of rooms is present and that the level remains traversable. The algorithm also inserts special rooms, such as boss rooms, obelisk rooms, puzzle rooms, etc., where necessary, ensuring variety and serving mechanics or storytelling.

The game's dungeon generation algorithm was implemented early in development, combining premade sets of rooms for each dungeon type in a randomized manner. This ensured that every dungeon appeared different each time a player restarted it. However, the algorithm occasionally generated identical rooms adjacent to one another, breaking immersion. One of the key points of interest for this research was to determine whether a dynamically changing dungeon layout influenced players' perception of the environment and overall experience.

Before the research, dungeon layout was randomized with every new run, including during level restarts. However, for the purposes of this research, it was decided to test difference in reaction to and experience with a static dungeon layout compared to a randomized layout. In the static version the room layout remained constant, while in the alternative build version layout changed every time player started or restarted the dungeon. This allowed to compare experiences particularly in terms of navigational memory, perceived complexity, and player satisfaction. Table 2 below showcases the steps taken in generating a dungeon in the static version of the Test Build.

Table 2. Dungeon generation in Willow Guard

No.	Step	Description
1	Room Selection	A script selects a room from a predefined list based on the area type
2	Exit Connection	Each selected room is assigned an exit direction (north, south, west, or east), ensuring connectivity with the next room.
3	Room Expansion	The process continues until a predetermined number of rooms have been placed.
4	Special Rooms	Key locations such as objective rooms and boss rooms are inserted at designated points in the level.

Randomized version of the Test Build was updated with a script further randomizing certain elements within the rooms in additional steps discussed in the following section.

5.2 In-room Randomization

The dungeon generation approach described in the previous subchapter has an obvious disadvantage. Regardless of the number of preexisting rooms, the script may occasionally select identical rooms and place them next to each other. A

simple solution for this issue could be adding a rule to the dungeon generation script excluding the possibility of connecting the same room prefabs to each other. However, the number of rooms in the dungeons on the later stages of the game increases significantly, and regardless of this rule identical rooms may still be generated close to each other on the critical path. Table 3 shows steps after the dungeon has been generated that randomize the contents of rooms in Test Build B.

Table 3. Randomization steps in Test Build B

No.	Step	Description
5	In-Room Randomization	Each room undergoes the process of randomization of its elements
5.1	Loot placement	Resources (gold, health bushes, crates, etc) are placed.
5.2	Environmental elements	Large decorative objects and environmental elements implemented through tilemaps and prefabs are randomized from a pool of options.

Both loot objects and environmental elements follow similar script. The number of elements necessary to place within the game space to keep the gameplay fair and balanced has been defined. This number of elements has been placed in the rooms of the Test Build A. Further, a larger pool of elements has been placed in the game space of Test Build B, with a script attached, which limited the number of elements spawned in each instance to the balanced number. Thus, both test builds were designed to have similar setup of collectible items, where Test Build A would have them in static spots, while Test Build B would randomize their placement.

This randomization was aiming to add variation and a sense of lived-in unpredictability through varied loot placement. From the narrative standpoint, loot containers could be removed or relocated, bushes might regrow in different patterns, and valuable items could shift locations between runs. This not only

supported the idea of potential replayability but also aimed to simulate a world that responds to time and player actions.

Figures 6 and 7 below present how collectible items have been placed in the level before and after the randomization script application. Figure 7 presents one of the possible combinations of collectibles placement reflecting actual amount of these items ultimately spawning in the dungeon.

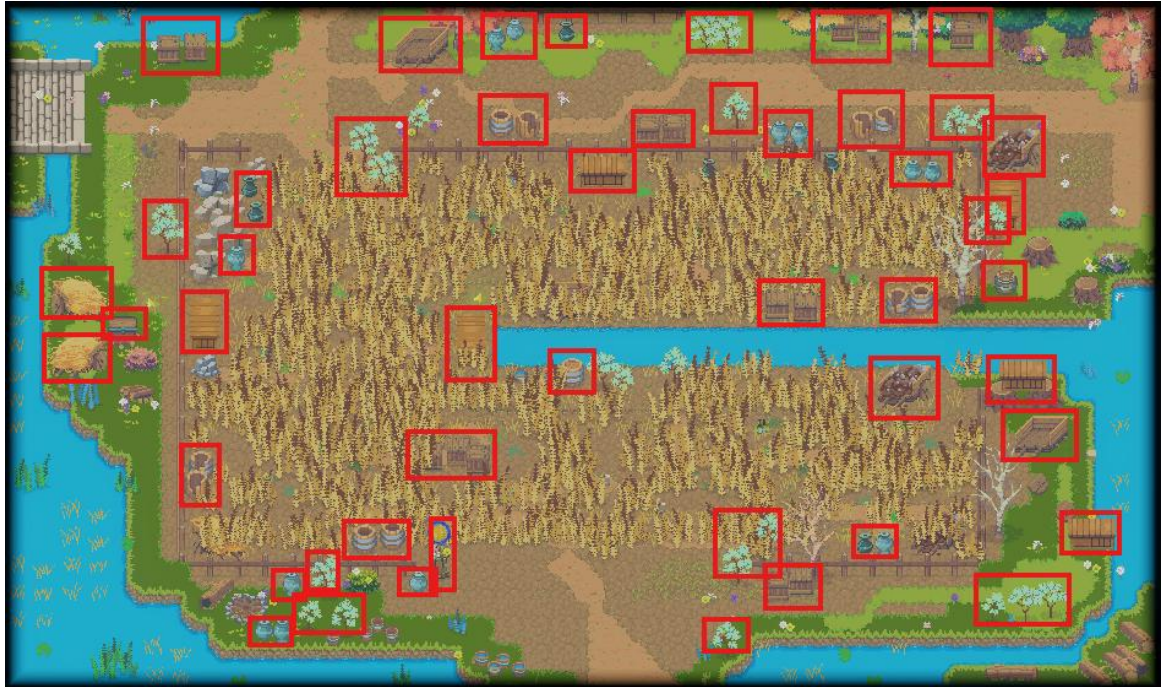


Figure 6. All collectible items in the level before randomization

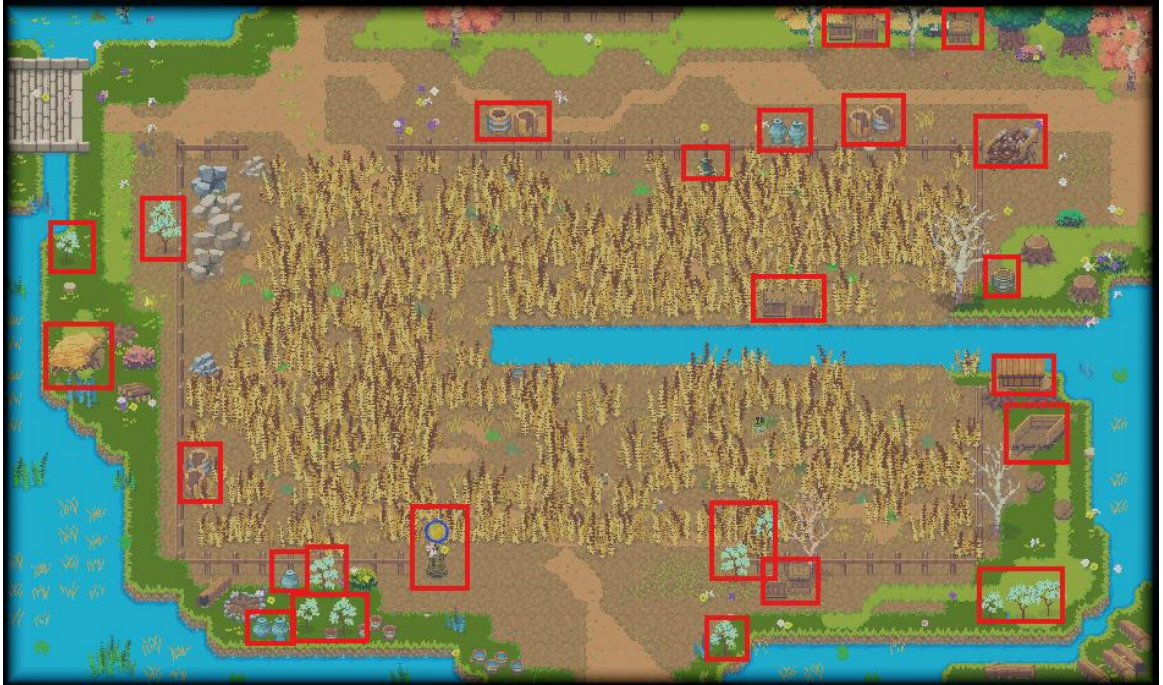


Figure 7. Instance of randomized collectible items

At the macro level, large outdoor areas such as crop fields and grassy plains were diversified in Test Build B. Field areas were edited to vary in crop density and include patches of either thick or sparse growth. Tilemaps used in creating levels contained two types of green tiles designating grass, both of which were used to create a more painterly and uneven natural landscape. The grass patch outlines were also diversified to make the areas look different. Additionally, footpaths were randomized using different terrain materials: stones, pebbles, dirt, or occasional overgrowth. This approach allowed the environment to shift subtly between playthroughs without impacting navigation or level flow. Figure 8 shows the same dungeon room as used in Figures 6 and 7, with different path and grass options, as well as randomized cornered environmental elements, and an alternative variant of the field crops density.



Figure 8. Instance of randomized environment elements and collectible items

On a smaller scale, each corner of a room (northeast, southeast, northwest, southwest) was populated with one of several hand-designed environmental elements. These included, but not limited to, wells, timber piles, stone barriers, forests, water banks, neighboring field corners, and other elements inherent to rural areas. The algorithm selected one variation per quadrant, resulting in a wide array of unique combinations.

6 RESEARCH FINDINGS

This chapter presents the results of the study comparing two distinct builds of the game: Build A, which utilized a non-randomized layout, and Build B, which incorporated randomized elements within the level design. The data was gathered through participant feedback, gameplay observations, and collecting subjective player experiences through interviews.

A total of 21 applicants filled out the tester application form, with 18 ultimately participating in the test, nine testers for Test Build A and nine for Test Build B. Each test build feedback form received nine responses, with three respondents having tested both test builds and responded to an additional comparison

feedback form. While the sample size of 18 participants is limited for drawing statistically significant conclusions, it remains valuable for generating insights within the scope of indie game development, where iterative feedback and qualitative evaluation are commonly applied.

6.1 Dungeon Layout

Player perception of the dungeon uniqueness differed notably between the two builds. In Test Build A, featuring static layout, repeated runs led to familiarity, predictability, and, in some cases, boredom. The latter observation was supported by statements about changing the strategy to only run through the rooms necessary to fulfill the objective after one or more deaths or restarts. A tester of both test builds commented the following on the dungeon layout:

“Playing the game with static layout felt more frustrating when dying since you had to start all over again, but with randomized layout it felt like a new level completely even with the same rooms.”

Test Build B layout, randomized to create a more dynamic experience through unpredictability in the initial conditions, proved to be more engaging and to have an increased replayability potential. Players reported that the randomized dungeon layout and logically different critical path made the level feel less predictable and more engaging, even though the core aesthetic elements did not change.

Although about half of respondents in both Test groups stated that starting dungeon layout had no impact on their (four in Test Group A and five in Test Group B), static layout had obvious negative effect on the other half of the Test group A, who stated it “felt repetitive and reduced [their] engagement”. In addition to this, two out of three testers who played both builds preferred Test build B:

“Playing the game with static layout felt more frustrating when dying since you had to start all over again, but with randomized layout it felt like a new level completely even with the same rooms.”

“When replaying one level multiple times, randomization increases engagement.”

This finding highlights the importance of layout randomness in enhancing player perception of level uniqueness and reducing frustration from repeating identical dungeons in cases of failure.

6.2 Collectible Items

Collecting loot items and resources was highlighted by both testing groups as having the highest effect on the motivation to explore levels in-depth (six respondents out of nine in Test group A and eight respondents out of nine in Test group B), and one of the highest ranked goals to play games in general. Naturally, it occurs from the fact that loot and resources are major factors of progression in many games.

Randomization of gathering spots of healing bushes, resource crates, and stamina boosts has not shown a notable impact on the players' behavior regardless of their self-assessed intrinsic motivation. Seven out of nine Test group A respondents have stated that loot placement had no or little influence on their exploration choices in the static dungeon. At the same time five out of nine respondents, or slightly over a half of Test Group B, stated to have explored more actively to find collectible items. This shift in randomized dungeon might be caused by players' noticing randomization, even if only subconsciously, and therefore seeking for loot items more actively.

One pattern prevalent among respondents of both groups indicated that frustration from dying and the necessity to restart impacted the willingness to explore and shifted the players' intrinsic motivation from exploration to completing the level objective.

6.3 Environmental Variations

Participant attention to the visual aspects of the test level was relatively comparable between the two test groups. It was evaluated using a 4-point scale, where "1" indicated minimal attention ("I play levels through as fast as possible") and "4" indicated maximum attention ("I spend extra time on the levels to

explore”). Eight out of nine respondents in Test group A and seven respondents in Test group B stated that that visual aspects are important for their gaming experience. None of the respondents selected the option of minimal attention paid to the visual aspects of the game.

The split test attempted to assess repetitiveness of the visual aspects in rooms as it appeared to the testers. It is worth mentioning again that the area selected for testing, namely fields, is homogeneous as it is in the real world, which decreases potential for variation. Regardless of this fact, in Test Build A, the visual appearance of the level was reported to feel repetitive by more respondents compared to the Test Build B. Particularly, from the results of personal interviews it became apparent that Test Build A players were fatigued by the environment slightly more, which resulted in lower numbers of replays beyond the three replays required by the test conditions.

Testers were asked to evaluate the number of unique rooms they encountered. Both test builds have been designed to generate a dungeon with seven unique rooms. Test Build A rooms are static, meaning testers cannot objectively experience more than this number of unique rooms. Test Build B also contains seven unique room prefabs; however, rooms in the Test Build B are implemented with the built-in randomization making up to dozens of combinations, however, with repeating elements. Testers in both groups provided the impression-based evaluation of the number of unique rooms they encountered. Test Build A was evaluated to contain an average of 6,66 rooms and Test Build B 8,33 unique rooms, which supports the hypothesis that players observe and notice variation even when it is only a part of the level that is modified.

6.4 Exploration

A notable difference emerged between the non-randomized Test Build A and randomized Test Build B versions regarding player exploration incentives and behaviors. In test Build A, players tended to explore thoroughly at first, but familiarity with the static layout led to reduced curiosity over time. Some testers

described their behavior in later runs as rushing through the rooms to the objective room, which indicated a decrease in exploration in subsequent runs.

In contrast, Test Build B encouraged continued exploration, as players were motivated to revisit areas in search of changes in layout itself, loot placements, and puzzle solutions. Players reported that the unpredictable nature of randomized elements helped sustain engagement and kept exploration relevant and battle encounters more engaging throughout the testing, which is supported by the following quotes:

Tester of Test Build A: "At first I was exploring with high precision, looting and killing and breaking everything I could find. But after dying and everything just being reset, I somewhat lost interest to exploring."

Tester of Test Build B: "If the primary objective was close I finished it first before exploring, otherwise exploration came first."

Tester of both test builds: "It was more exciting to explore dungeon rooms when they were randomized since you never really knew what you'll face in the next room."

It is worth noting additionally that increased difficulty of the game for some players can cause frustration and decrease their motivation to complete other than main objective tasks, which observation was pointed out by testers regardless of their Test group.

6.5 Key Findings

Randomization in Test Build B resulted in sustained exploration motivation across multiple playthroughs, compared to the decreased exploration observed in Test Build A over time.

Randomization of the dungeon layout in Build B led to a greater sense of uniqueness and, therefore, sustained interest, it reduced predictability in each playthrough, contributing to a more dynamic player experience, strategy adjustments, and unexpected outcomes.

Randomization in loot placement has not proved to be a factor contributing significantly to the players' experience. Certain technical constraints also evolved during the implementation of the Test Build B with randomization, that are outlined in the Discussion section, that suggest considering necessity of loot randomization.

The results of this study clearly indicate that randomization in linear level design can noticeably enhance player engagement, exploration, and replay value. While Test Build A, with its fixed layout and contents, resulted in predictable patterns and decreased motivation over time, Test Build B's randomized elements kept the experience dynamic and fresh, encouraging players to explore, adapt, and engage with the game repeatedly. These findings support the hypothesis that randomization can be an effective tool in reducing monotony and improving the overall player experience in linear level designs.

Finally, player feedback regarding memorable moments and frustrations was analyzed. In Test Build A, players frequently mentioned predictability as a source of frustration. After a few runs, they felt they had mastered the layout and mechanics, leading to a sense of boredom and disengagement. Test Build B also averaged more replays, due to this, among other reasons. Furthermore, issues such as backtracking and the lack of dynamic changes contributed to negative experiences.

In contrast, Test Build B generated more memorable moments. Players reported moments of surprise when they encountered random elements that altered their strategies, such as unexpected enemy encounters or changes in critical path. This sense of novelty and unpredictability contributed to a more dynamic experience, while frustrations related to repetitive gameplay were notably reduced.

Testers comparing Test build A and Test build B showed the following pattern in assessing impact of the dungeon layout. Testers reported either negative or no effect of the static dungeon on their engagement with the level and game over

time, strategy to collect loot, appeal for replayability in both cases of recurring dungeon or death. Simultaneously, they reported positive or no effect made to the same aspects (engagement over time, strategy to collect loot, appeal for replayability in cases of recurring dungeon or death) by the dungeon layout randomness.

These key findings signify a positive answer to both the main and the secondary research questions: Yes, players are able to perceive increased variation of the levels, and the increased variation indeed improves players' impression of the game in general, its replayability value, their curiosity and long-term immersion.

7 DISCUSSION AND ANALYSIS

The following chapter discusses the results of the qualitative A/B test performed to assess the impact of randomization on player experience, compares these results to the ones from relevant previous research, and provides recommendations on further development of Willow Guard.

7.1 Comparing Findings to Previous Research

While previous research offered conflicting conclusions, where Conner et al. (2017) stated handcrafted levels to possess higher immersive potential, and Rodrigues and Brancher (2017) reported no significant difference between static and procedurally generated environments, this study suggests that a middle ground may offer the best of both worlds.

In contrast to Rodrigues and Brancher's (2017) findings, where procedural variation in a mathematical game had little effect on player engagement, the randomized level design of Test Build B in Willow Guard resulted in higher replay motivation and sustained player interest over time. This outcome likely reflects genre-specific factors, for action role-play games with its exploratory and deck-building elements appear to benefit more substantially from spatial variation and unpredictability, compared to abstract or simple games.

Furthermore, the results complement Conner et al. (2017) emphasis on the immersive qualities of handcrafted environments. Although Test Build B employed randomization, it did so within the boundaries of a deliberately structured design. By preserving narrative coherence while allowing for variation in layout and collectible items placement, the build maintained a sense of intentionality supporting Bastos et al. (2017) concept of “credible” design that aligns visual and mechanical elements to foster immersion.

However, not all randomized components contributed equally to player experience. The randomized loot placement, for example, did not yield notable changes in perceived variation, nor increased engagement, echoing Bastos et al. (2017) notion that believability and relevance are more important than sheer variation. This nuance emphasizes the importance of purposeful design even within systems that embrace randomness.

Overall, this research demonstrated that procedurally randomized elements, when integrated thoughtfully into otherwise linear environments, can enhance immersion, exploration, and replayability. This conclusion bridges the gap between past findings by validating the value of controlled variability and reaffirming the importance of coherent, player-relevant design choices in immersive gameplay.

7.2 Recommendations for Level Design in Future Development of Willow Guard

Based on the findings of this study, it is evident that players do perceive and respond positively to variation in level design. The introduction of randomization in layout and environment notably increased engagement, curiosity, and replayability. However, randomized loot placement did not yield a comparable impact and may not justify the additional development complexity under current constraints. Therefore, future efforts should prioritize spatial and structural variation over randomizing auxiliary content.

Each level template can then be divided into spatial sectors: four corner quadrants (northeast, southeast, northwest, southwest), four exits (north, south, east, west), and a central area, resulting in nine definable segments. For each of these nine sectors, several combinations of in-game objects should be created and placed in the sector prefab as alternative options. These variants must adhere to biome requirements and reinforce the environmental narrative of the level. The script used for implementation of Test Build B can then be used to randomly select one prefab per sector during level generation.

To effectively implement this, it is recommended that future levels are constructed from a modular set of templates and sector-based prefabs. Specifically, designers can create prefabs with tens of options before performing the level design itself, and leaving three to four alternatives suitable for the respective sector. Level designers can optionally edit these remaining alternatives to make them visually and thematically distinct regardless of them being reused. These prefabs might emphasize different aesthetic elements, such as water features, ruins, dense forests, or expansive plains, aligning with the environmental storytelling of the game.

Assuming a base of three level templates and three prefab options per each of the nine randomized sectors, the total number of potential unique combinations skyrockets at a low computational power. The prefabs can also be individually modified to better fit in the template to further differentiate the similar elements.

By leveraging this modular and randomized approach, future development of Willow Guard content can increase long-term engagement at the same time achieving greater efficiency. These are two decisive factors for the indie developers' decision making as they often strive to align with production constraints and maintain artistic direction. This recommendation represents a practical and scalable strategy for integrating controlled randomness into linear level design.

8 CONCLUSION

This research demonstrates that introducing controlled randomization into linear level design can significantly enhance player experience by increasing engagement, replayability, and immersion. While handcrafted content remains essential for narrative and structure, variation within that structure helps maintain player interest across multiple playthroughs. The findings support a balanced approach, where deliberate design is augmented by randomized elements to create a more dynamic and personalized experience.

Notably, not all types of variation proved equally impactful. Randomized layouts and environmental variation contributed meaningfully to players' sense of exploration and strategic thinking. In contrast, loot placement variation had a negligible effect on player experience and may not justify the design effort or time contribution required to implement it. These insights allow to answer if players notice and value the increased variety within the levels: most elements are indeed positively perceived by the players; however, some of the test subject elements (such as loot placement) were not perceived enough to consider their randomization in future development. These results help prioritize where variation matters most, enabling developers to focus resources efficiently.

The use of two distinct builds to test opposing hypotheses, one static and one randomized, provided a clear comparison that strengthened the conclusions. Two contrasting hypotheses have been formulated by the development team and researcher, stating that players would or would not notice increased variation. Effectively, both hypotheses turned out to be correct in relation to a specific aspect of randomization. Players proved to have perceived increased variation in environment elements, but not collective items placement. Even with a limited sample size, the results offer actionable recommendations for indie developers seeking scalable ways to enrich their games without drastically expanding scope. Observed patterns in player behavior and feedback reveal that even subtle variations can have a noticeable cumulative effect.

While the findings offer valuable insights, this research is subject to certain constraints that must be acknowledged. The small sample size limits the statistical generalizability of the results, and the study's reliance on subjective player feedback introduces potential biases. Additionally, the limited duration of gameplay sessions may not capture long-term engagement patterns or emergent behaviors over extended play. Despite these constraints, the internal validity of the study remains strong due to the controlled comparison between two well-defined builds. Still, further studies with larger and more diverse participant groups are needed to confirm and expand upon these findings. Future research could also explore genre-specific responses to level variation, long-term retention effects, and the impact of other types of randomization such as narrative branching or adaptive enemy behavior.

Future level design in Willow Guard and similar projects can benefit from modular randomization strategies. Building a small set of distinct templates and combining them with interchangeable, biome-relevant elements allows for a high number of unique configurations while preserving design coherence. This approach provides both variety and efficiency, supporting sustainable development practices in small teams aiming to deliver engaging, replayable experiences.

REFERENCES

Bastos, A., Gomes, R.F., Costa dos Santos, C. & Maia J.G.R. 2017. Assessing the Experience of Immersion in Electronic Games. Conference paper. November 2017. 19th Symposium on Virtual and Augmented Reality. Available at: https://www.researchgate.net/publication/321176135_Assessing_the_Experience_of_Immersion_in_Electronic_Games [Accessed 14 April 2025].

Bycer, J. 2018a. Replayability in Game Design. Blog. May 18 2018. Available at: <https://medium.com/super-jump/replayability-in-game-design-798fbb91a726> [Accessed 14 April 2025].

Bycer, 2018b. Level Design vs. Environment Design. Blog. April 16 2018. Available at: <https://medium.com/super-jump/level-design-vs-environmental-design-b8d19992924e> [Accessed 14 April 2025].

Connor, A.M., Greig, T.J., & Kruse, A.M. 2017. Evaluating the Impact of Procedurally Generated Content on Game Immersion. *The Computer Games Journal*, 6(4), 209-225. Article. Available at: https://www.researchgate.net/publication/319173732_Evaluating_the_Impact_of_Procedurally_Generated_Content_on_Game_Immersion [Accessed 12 March 2025].

DeLally, M. 2020. Level Progression and Pacing in Puzzle Games. Blog. August 18 2018. Available at: <https://medium.com/@mdelally/level-progression-and-pacing-in-puzzle-games-eb9e6a97e571> [Accessed 14 April 2025].

Ginting, A.K., Sari, K., Fadhilah, C. Hartama, D., Yusra R.N. & Zarlis, M. 2019. Application of the Perlin Noise Algorithm as a Track Generator in the Endless Runner Genre Game. *Journal of Physics Conference Series*, 1255(1). Available at: https://www.researchgate.net/publication/335660507_Application_of_the_Perlin_

Noise_Algorithm_as_a_Track_Generator_in_the_Endless_Runner_Genre_Game [Accessed 14 April 2025].

How Does Procedural Generation Work? 2020. DigiDigger. Video clip. Available at: <https://www.youtube.com/watch?v=-POwgollFeY> [Accessed 14 April 2025].

Kekki, J. 2024. A/B Testing: How (not) to Visualize A/B Tests. Webpage. Available at: <https://metacoregames.com/news/a/b-testing-how-not-to-visualize-a/b-tests>

Kremens, R. 2009. Level Design: Concept, Theory, and Practice. Boca Raton: CRC Press.

Miszczyk, M. 2017. Arcanum: Of Steamworks and Magick Obscura. Blog. Available at: <http://www.hardcoregaming101.net/arcanum-of-steamworks-and-magick-obscura/> [Accessed 14 April 2025].

Laitaruk, I.F. & Hryshanovych, T.O. 2024. Overview of modern algorithms for world procedural generation in computer games. *CEUR Workshop Proceedings*, 3917, 152-163. Available at: <https://ceur-ws.org/Vol-3917/paper21.pdf> [Accessed 14 April 2025].

Porokh, A. 2023. Introduction to game design Environment and initiating your journey. Blog. Available at: <https://kevurugames.com/blog/understanding-and-initiating-game-environment-design/> [Accessed 14 April 2025].

Roberts, S. 2015. John Horton Conway: the world's most charismatic mathematician. *Guardian*, 23 July 2015. Available at: <https://www.theguardian.com/science/2015/jul/23/john-horton-conway-the-most-charismatic-mathematician-in-the-world> [Accessed 14 April 2025].

Rodriguez, L. & Brancher J. 2019. The Impact of Procedural Level Generation on Players' Experiences and In-game Behavior. Conference paper: Brazilian

Symposium on Computer Games and Digital Entertainment (SBGames 2019).

Available at:

https://www.researchgate.net/publication/337032924_The_Impact_of_Procedural_Level_Generation_on_Players'_Experiences_and_In-game_Behavior [Accessed 14 April 2025].

Sainio, N. 2023. Terrain Generation Algorithms. MS Thesis. Available at:

<https://trepo.tuni.fi/bitstream/handle/10024/147549/SainioNiko.pdf;jsessionid=1398CC0B394F174DF239E28284DD5DBA?sequence=2> [Accessed 14 April 2025].

Salmond, M. 2021. Video Game Level Design: How to Create Video Games with Emotion, Interaction, and Engagement. London: Bloomsbury Publishing Plc

Sampaio, P., Baffa, A.C.E., Feijo, B. & Lana, M. 2017. A Fast Approach for Automatic Generation of Populated Maps with Seed and Difficulty Control. Conference paper. November 2017. 16th Brazilian Symposium on Computer Games and Digital Entertainment. Available at:

https://www.researchgate.net/publication/326153282_A_Fast_Approach_for_Automatic_Generation_of_Populated_Maps_with_Seed_and_Difficulty_Control [Accessed 14 April 2025].

Stipcevic, M. & Koc, C.K. 2014. True Random Number Generators. In Koc C.K. (ed.) Open Problems in Mathematics and Computational Science: 275-315.

Available at:

https://www.researchgate.net/publication/299824248_True_Random_Number_Generators [Accessed 14 April 2025].

Totten, C. 2016. Level Design: Processes and Experiences. Boca Raton: CRC Press.

Totten, C. 2019 Architectural Approach to Level Design. 2nd ed. Boca Raton: CRC Press.

Wilson. J. 2018. Level Design: Understanding a level. Blog. Available at: <https://www.gamedeveloper.com/design/level-design-understanding-a-level> [Accessed 14 April 2025].

von Neumann, J. 1951. Various Techniques Used in Connection with Random Digits. *National Bureau of Standards Applied Math Series*, 12 (1951), 36-38. Available at:

https://mcnp.lanl.gov/pdf_files/InBook_Computing_1961_Neumann_JohnVonNeumannCollectedWorks_VariousTechniquesUsedinConnectionwithRandomDigits.pdf [Accessed 14 April 2025].

Willow Guard Playtest Feedback Form

Your gaming preferences

1. What is your main goal while playing a dungeon (in general)? – multiple choice question
 - Combat and battles
 - Exploration and discovering new and/or hidden areas
 - Collecting loot and resources
 - Solving puzzles and interacting with environment
 - Engaging with story and lore
 - Strategy and gear
 - Speedrunning
 - Other

2. If you chose Other in the previous question, please specify – question requiring long answer.

3. How do you typically approach a new dungeon in a game? – multiple choice question
 - I explore thoroughly before progressing
 - I focus on the main objective but check the side paths
 - I move towards the main objective avoiding distractions
 - I complete the objective and explore level afterwards

4. What motivates you to explore a dungeon in-depth?
 - Finding new loot and resources
 - Uncovering lore and story elements
 - Seeking combat encounters or challenges
 - Enjoying the environment and setting
 - Completing objectives and/or achievements
 - Other

5. If you chose Other in the previous question, please specify – question requiring long answer.

6. How much attention do you usually pay to the visual aspects of dungeons?
 - Scale
 - 1 - I play levels through as fast as possible (minimal attention paid)
 - 4 - I spend extra time on the levels to explore (maximum attention)

Testing Willow Guard

7. How many times did you start the level anew (won / died / restarted)?
 - 1 time (I spent all the playtest time playing without dying or restarting)
 - 2-3 times
 - 4-6 times
 - 7 or more

8. Did you feel like you were improving your skills with each attempt? – short answer

9. How much time did you spend exploring the rooms?
 - Minimal exploration, I focused on completing the objective
 - Some exploration, but mostly followed the objective
 - Some exploration, occasionally strayed from the main path for loot or curiosity
 - Extensive exploration, I searched for hidden areas and tried to collect all loot

10. What was your approach to moving through the dungeon?
 - Throughout the playthrough, I wanted to explore as many rooms as I could
 - I explored all rooms in the beginning, and moved to the objective room after a couple of playthroughs
 - I only explored the rooms I thought were necessary for the objective
 - I aimed at completing the main objective in the beginning, and explored more after a couple of playthroughs
 - Other

11. If you chose Other in the previous question, please specify – question requiring long answer.

12. What was your approach to collecting loot during the playtest?
 - I searched every possible location
 - I prioritized obvious loot spots but didn't search extensively
 - I collected loot only when it was in my direct path
 - I didn't actively search for loot
 - Other

13. If you chose Other in the previous question, please specify – question requiring long answer.

14. How did the loot placement influence your exploration choices?

1 Loot placement had no influence at all

4 I explored more actively searching for loot

Willow Guard level design

15. How did the starting layout affect your experience?

15a. How did the static starting layout affect your experience? (Question specific to static dungeon layout)

I didn't notice the starting layout was static

It didn't strongly impact my experience

It felt repetitive and reduced my engagement

It gave me a sense of familiarity and mastery

15b. How did the changing starting layout affect your experience? (Q specific to randomized dungeon layout)

I didn't notice the starting layout was static

It didn't strongly impact my experience

It felt confusing and reduced my engagement

It gave me a sense of novelty and variety

16. How many unique rooms have you encountered during the whole playthrough (impression-based estimate)?

17. Did certain visual aspects of the rooms feel repetitive to you?

18. Did anything about the level design of the dungeon stand out to you?

19. Please leave your comments in free form as well