Balancing Randomness and Predictability in Games

Manuel Correia
Instituto Superior Técnico (of University of Lisbon)
Lisbon, Portugal
manuelmpcorreia@tecnico.ulisboa.pt

***Abstract*—Randomness is a useful tool in game design as a source of excitement and novelty. Nevertheless, it is often troublesome to balance. Overdo it, and the player will feel no control. Else, if it is minimal, it may be redundant. This work hypothesizes that a dynamic adjustment of randomness can positively impact the player experience (PX). First, we analyze previous works on the human perception of randomness and how it shapes game design. Several use cases are reviewed to understand how it informs the PX, and how to balance that. There is also an investigation on how to evaluate the potential changes to these aspects. Second, we formulate a model for this adjustment based on the player and intervening entities' resources throughout a playthrough. The selected case study was boss battles, challenges often difficult to balance given player disparities. With Enter the Gungeon as a testbed game, the decision-making behavior and respective action functionality involving random number generation (RNG) of the adversary were modified. Participants played against three versions: one unpredictable, one predictable, and one dynamic; then reported on their experiences. On the reported aspects, it was shown that players felt more competent with predictable behavior, significantly more than with an unpredictable one. Similarly, the dynamic behavior made participants feel more capable of growth. Moreover, there were indications that players who perceived themselves in control, according to a Locus of Control assessment, preferred predictable experiences. On the other hand, this mindset seemed connected to more effort put into their performance.**

***Index Terms*—Randomness, Dynamic Adjustment, Balancing, Virtual Games**

I. Introduction

A. Motivation

Throughout video games, chance plays a pivotal role in defining the user's experience. Unpredictability can both create replayability and excitement when appropriately employed. Likewise, removing chance in games that hinge on deterministic behavior is essential. However, some scenarios are not clear cut. For instance, when randomness plays a part in the level of the challenge the player faces. Unpredictability can be a source of difficulty beyond the player's control and, thus, a point of frustration. Entirely predictable behavior, on the flip side, may feel repetitive. And, different players, based on their expertise, have different definitions of what constitutes an appropriate hurdle. For these reasons, reaching a static implementation that suits a broad player base is not simple or even possible.

B. Problem

As video games have become an ever more ubiquitous form of entertainment, balancing the player experience (PX) has, too, grown in complexity.

As a game mechanic, randomness can impact several aspects of the PX, such as excitement and difficulty. With large or growing player bases, a static solution will struggle by not suiting all players and, in trying to, may please none. Beyond this, randomness can create highly variable gameplay experiences that incentivize replayability. As a result of this design, players grow to play significantly differently throughout their time with the game, invalidating existing balance. Here, the consequence of static solutions is repetitiveness, which is antithetical to the overall design.

As such, we will address the problem of effectively balancing randomness to create a PX uniformly satisfactory for a growing audience, with increasingly diverse players.

C. Hypothesis

The player resources can be indicative of performance, which can inform the dynamic adjustment of randomness. This process can take the PX from pure chaos to rigid patterns, potentially improving it.

When facing a more powerful adversarial entity in the game, a player bridges the power gap through strategy and mastery of the controls. These challenges, referred to as 'boss fights,' are used as a test to conclude a level. Throughout these confrontations, either the player or enemy will get the upper hand, and a comparison of the player resources and the enemy resources can determine which (e.g.: health percentage). That can then be the basis for adjusting the behavior of the player's adversary. If the player has substantially more resources than the enemy, the adversary's behavior can become unpredictable, limiting the player's ability to strategize. Or, were the enemy to be the one with the most resources, they would adjust to follow clearer patterns, giving the player more breathing room. This responsiveness should result in changes both to the player's reported perception of the gameplay and the logged metrics during the playthrough. This data should, in turn, allow for an evaluation of the impact of the manipulation of randomness.

With this, we hypothesize that by dynamically adjusting the predictability of the adversary's behavior based on the resources of both entities in real-time, the PX can be improved by better suiting each player.

II. Related Work

Randomness is a versatile and useful tool for game design, but the concept's complexity means it is not trivial to implement.

Looking into randomness as a concept, based on its application, it can be divided into to categories based on the point in time at which it impacts the experience [1]:

- Output Randomness - "noise injected between the player's decision and the outcome" (e.g.: rolling a die to determine if an attack hits the target);
- Input Randomness - "informs the player before he makes his decision" (e.g.: procedural level generation).

And, in either scenario, especially in games, its application would classify it as 'controlled randomness'. Its implementation is most often manipulated in some respect, particularly to avoid sequences of repeated results or other such patterns. Though these are common within randomly generated sequences, humans do not perceive them as such.

Through my analysis of existing work on the perception of randomness [1], it is clear that human perception is skewed. To start, our evolution as a species led to the development of mechanisms meant to aid in our survival. However, applying these innate capabilities to problems they are not suited for can instead be to our detriment. To survive, we became better at discerning patterns and creating connections, for that reason, randomness confounds us as it defies those expectations. In looking for patterns where by definition there are none, or, in the case of pseudo-randomness, they are too complex to discern, humans can make bad decisions [2].

Considering the above, one might think that randomness would be something to eliminate from experiences such as games, but it has long proven its value. Precisely because it defies our expectations, the introduction of randomness in games, from the uncertain outcomes of our actions to the construction of similar but unpredictable circumstances, creates novelty and excitement [3]. Despite our instincts that fight for survival, we crave experiences with risk, as knowing every outcome leads to monotony. A game like chess does not introduce randomness through any mechanics, and yet, it can be very exciting as we can only try to predict what the opponent will do next.

It is trivial to say that excitement and novelty are common motivations for engaging in an activity. On the other hand, from the work we have referenced [4], it has been shown that feeling in control, being capable of enacting change, more precisely, the change we intend, upon an environment, are motivators as well. Choosing to play a given game, much like a sport or other hobby, can be tied to our feelings of competence in that task and our perception of improvement [4]. When failing repeatedly at something we become frustrated, all the more so if, with each attempt, there is no sense of progression, of improvement. For this reason, randomness is tough to balance, as a poor implementation can mean the player feels their actions are inconsequential.

As discussed, it is a tool that can create novel experiences from the same building blocks (procedural generation), and perhaps, from a player's perspective, this is the least frustrating approach. Though their starting circumstances are different each time, the outcome is tangibly a result of their actions. Conversely, using it to make the outcomes of player actions uncertain is much more nuanced. A player can accept that at times they may be unlucky, but not feeling they lack agency. That could be why, much like we cannot predict with certainty the actions of an opponent in chess, the use of randomness to shape the behavior of adversaries not only is exciting but also is deemed fair in the player's eyes.

To evaluate a game's use of randomness, it is the player experience that must be analyzed. After all, games, digital or otherwise, are experiences crafted for the player's enjoyment. Agency is perhaps the most relevant component of that experience when evaluating the implementation of randomness. In many mediums, the consumer is a passive agent, their only role to observe, but games are interactive experiences, and though in many cases the overall narrative is unchanged by player action, there is much enjoyment in furthering it. People enjoy books and movies, regardless of the lack of interactivity as the locus of control was never in question. When playing a game, the locus of control should be internal, lest players be victims rather than participants.

In addition, different individuals will have varying mindsets, meaning the individual's outlook, just as the experience, must be considered. To assess each person's perception of their own agency, Rotter's Scale for Locus of Control (LoC) classifies people between an external and internal LoC. Those with the external mindset believe that outside influences, chance, or fate, have the bigger influence on the course of their lives. Alternatively, an internal mindset corresponds to the view that one's actions are the prevailing force.

In conclusion, randomness is an important game mechanic when used effectively. Controlled randomness, for one, can increase replayability, but the implementation of output randomness requires complex balancing. As has been done with difficulty, one answer to this problem might lie in its dynamic adjustment. Not all players are alike, as mentioned, and so, testing and balancing while in development can only accomplish so much. Creating systems that alter the experience is effective in improving player satisfaction [5]. Specifically, an implementation that alters how random the experience is, based on player performance [6].

III. Case Study

A. Testbed Game

For this work, we will be modifying the game Enter the Gungeon [2] (ETG). This 'bullet hell' rogue-like has been highly rated by players, having an "Overwhelmingly Positive" rating on the PC platform Steam. Its gameplay experience is well established and does not require further validation. This sub-section aims to provide an understanding of the game's relevant aspects for this work.

[1] K. Burgun, Randomness and Game Design. Last accessed 6 Nov 2020 - https://www.gamasutra.com/blogs/KeithBurgun/20141015/227740/

[2] Dodge Roll, "Enter The Gungeon," Devolver Digital, 2016 Available: https://dodgeroll.com/gungeon/

Fig. 1. Gatling Gull Boss Encounter (Experiment Version) - Enter the Gungeon

Players explore procedural generated floors that culminate on boss encounters. These confrontations pit players against enemies with intricate skills using projectiles. Likewise, the player character's main resource are the weapons collected during play, each one unique. To overcome these challenges, they must also make use of two other central mechanics: dodging and blanks. The former, removes collision with all projectiles while the character rolls forward. The latter, creates a small explosion which clears all enemy projectiles on-screen. With a substantially lower health point count, player's must strategize and carefully use the tools available to surmount these fights.

Developed in Unity 3D, this game's code can be decompiled to help with developing modifications. Alongside this, it has an active modding community which has created several tools to help these endeavours.

B. Boss Behavior Model

The Boss NPC studied will, until the encounter concludes, cycle through the following model:

1) Evaluate its present resources;
2) Evaluate the players resources;
3) Compare the two evaluations to decide on one of three scenarios:
    a) It is loosing to the player;
    b) It and the player are on equal-footing;
    c) It is winning.
4) Based on the determined scenario, opt, respectively, for one of the following behaviors:
    a) Chaotic - the next action is chosen arbitrarily, with all choices available equally likely to be picked (only the action just executed is excluded);
    b) Balanced - the next action is chosen randomly, with each having an assigned likelihood, making some more frequent than others;
    c) Sequential - a predetermined sequence of actions exists which the Boss will follow, thus, in this mode, the behavior is entirely deterministic.
5) Based on the selected behavior, decide on an action;
6) Once the action is concluded, repeat from 1. until the fight concludes.

Beyond this modified decision-behavior, the use of random number generation [3] (RNG) within the actions available to the boss, prompted their alteration. These actions were changed to adjust the magnitude of the randomness within their calculations based on the boss variant and current behavior profile.

C. Implementation

To test my hypothesis, the testbed game needed to be modified on three main fronts: the player character, the flow and environment, and boss behavior.

1) Player Character: Players interact with the game through a character with resources and abilities which shape those interactions. In this section, it is explained what was available in ETG and how the implemented solution sought to constrain participants, given the wide choice available, in order to construct a consistent experiment scenario across players.

a) Game State: For rogue-likes, a core-tenet of the experience is randomization. Player characters then serve to give some control back, giving players the opportunity to somewhat shape their starting conditions. When starting a new playthrough in ETG, players can choose from several characters which vary mostly in starting equipment. However, across all, some fundamental aspects are consistent:

- the ability to dodge;
- a starting weapon with infinite ammunition;
- one passive item.

b) Implemented Solution: With too much inherent variability in the original design, the decision regarding the player character was that it ought to be standardized such that its resources were adequate for participants to engage with the experiment. As such, all participants were made to play the 'Pilot'. This is the first character any player of ETG is introduced to. He has very basic gear, the abilities of which would not interfere with the experiment. Two changes were made to the character's resources. His health was increased from six to ten health points, giving all players more of a chance to experience boss behavior. And, the magazine capacity of his gun was increased to reduce the frequency of reloads, which brake the pace of combat.

2) Experiment Flow and Environment: To accommodate the experiment design, several aspects of the game's flow needed changing. The two main objectives were to guarantee players of all experience levels would be given the necessary knowledge to participate, and to create an environment in which noise would not be introduced into the collected data through systems or mechanics extraneous to the present work.

a) Tutorial: First, as not all participants would necessarily be familiar with the game, it was important that the tutorial always be the starting point. This would guarantee basic knowledge of the essential mechanics, from shooting

[3] Random Number Generation (RNG) - a process by which a random number is generated, often used in games as means to randomize a parameter of a mechanic.

to dodging. Meeting this requirement meant that all entry-points of the game needed to be redirected to the tutorial, and that exit-points of the tutorial must lead to the test level. Otherwise, the base game's tutorial taught the player all required knowledge to participate in this work's experiment.

b) Test Level: Second, ETG's levels are procedurally generated and, as discussed, all this variation is not conducive to this work. Thus a test level had to be built in order to accommodate the experiment design. The alterations would have to allow for a pre-designed layout and facilitate the replaying of encounters by participants.

Thus, the solution was the creation of a custom level with four rooms. The player would begin the level in a central room that connected to three others (see Fig. 2), each of which contained a variation of the boss encounter. To avoid the introduction of bias through the sequence of the battles, which variant was in each room was set to one of three different sequences at the start of the game, using Latin Square (Table I). During the experiment the rooms were labeled from A to B anti-clockwise, and participants instructed to face the encounters in alphabetical-order. They were also allowed to replay each as many times as they wished by restarting the level, which would reset all encounters.

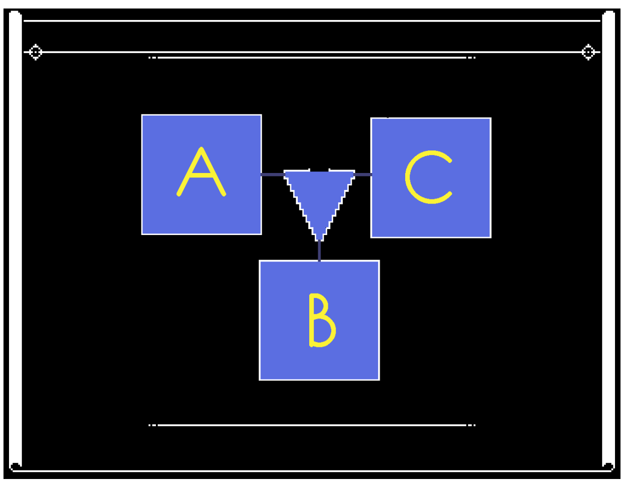

Fig. 2. Test Floor Layout

Sequence ID	A	B	C
0	Sequential	Chaotic	Dynamic
1	Chaotic	Sequential	Dynamic
2	Dynamic	Chaotic	Sequential

TABLE I
LATIN SQUARE BASED FIGHT SEQUENCES

3) Boss Encounters: As described (Sec. III-A), bosses in ETG use RNG within their action decision-process. To add to that, the actions themselves often randomize elements of their functionality. For this work, both of these aspects required modification to manipulate the magnitude of the randomness.

Originally, the selected boss - the 'Gatling Gull' - had a repertoire of six actions available, from which to chose (see Table II). [4]

Action Name	Description	Weight	Condition(s)
Walk and Shoot	The boss follows the player while shooting a random spray of projectiles.	3	None
Fan Spray	The boss, while stationary, shoots a wide spray of rapid bullets towards the player.	2	None
Big Shot	The boss shoots a large slow bullet towards the player, upon impact it explodes into multiple projectiles which go in all directions.	1.5	None
Waves	The boss, while stationary, shoots two waves of bullets towards the player with a small interval between them.	1.5	None
Leap	The boss jumps to one of several pre-assigned spots in the room.	1.5	Available leap-spots within the room.
Rockets	The boss executes Leap and then shoots rockets towards the ceiling which come down in spots marked by red cross-airs on the ground and explode.	3	Can only execute after Leap.

TABLE II
BASE GATLING GULL ACTIONS AND RESPECTIVE DESCRIPTION

Considering the existing implementation of boss behaviors, a new solution was necessary that allowed for a boss to pivot based on the criteria outlined in the model (see Sec. III-B). Nevertheless, the existing functionality was useful in creating the chaotic pattern and within the dynamic pattern for both the chaotic and middle profiles.

a) Chaotic: For this variant it was possible to reuse the game's original implementation. Its aim is to provide a contrasting implementation which is highly unpredictable and entirely static. This means that by altering the weights for all actions to 1, this could be accomplished. For the sake of player perception, the limitation that the same action could not happen twice in a row was added.

b) Sequential: In contrast, the sequential variant required a completely custom solution for its decision-making. This variant too was included for contrast, providing a entirely predictable experience. To that end, a sequence of actions was designed through which the boss would loop. The sequence [5] used was designed based on the original implementation of the encounter (see Table II).

c) Dynamic: At last, this works cornerstone, the dynamic variant, combined the previously outlined variants, with an intermediary behavior profile to bridge the two 'extremes'.

As laid out in the model (see Sec. III-B), this pattern assesses the player's and boss' resources to determine who is winning and choose a behavior profile. Multiple options were considered for this evaluation, in the end a naive approach

[4]The boss has one more action available, a melee attack, but this one cannot be selected in the above process and is instead executed reflexively when a player comes into range.

[5]Sequential behavior followed the sequence: 0, 2, 3, 0, 5. (Indexes of actions in the order listed on Table II)

was chosen: focusing only on health percentages. If the two's health percentages differed more than 7.5% then the boss would operate in one of the two 'extremes'. In sequence, if its health was highest, or chaotically if it was lowest. That 15% window would have the boss behaving with a 'moderate' level of randomness, with most actions using their original weights. Both 'Fan Spray' and 'Leap' were excluded, meaning their weights were 0, regardless of behavior pattern.

On the one hand, for both the chaotic and sequential profiles, the boss behavior followed the outlined for the respective standalone patterns. On the other hand, the 'moderate' profile uses the original implementation as a middle ground. Given that it utilizes random selection with assigned weights for each action, it achieves a semblance of predictability without being so.

d) Additional Alterations: Having set out how the behaviors operate, it is necessary to tackle the actions they employ. Of the four used within the experiment, two actions use RNG within their execution: 'Walk and Shoot' and 'Rockets'. For both of these actions, modifications were made to their processes, such that the actions' outcomes were more predictable in the sequential variant/profile and, conversely, more unpredictable for the chaotic variant/profile.

IV. Evaluation

In the context of this thesis, two experimental phases were carried out. A preliminary phase, which served to balance the experiment scenario and test its execution. And a main phase, to investigate the validity of the hypothesis.

A. Procedure

For both phases of research, tests were conducted remotely and followed the same procedure. All participants were provided with a link to an online form which would guide them through the experiment. It was structured as such:

1) Disclaimers - information on how collected data would be treated;
2) Demographic Data - Age, Gender, Player profile;
3) Mod Installation - a step by step guide, easily understood, to install the modifications to the game;
4) Experiment Procedure - a walkthrough of the procedure to carry out the experiment, from tutorial to facing each of the boss variants;
5) Experience Evaluation - an assessment of each boss encounter and a comparison of the variants.

At the end of the experiment, before the submission of the form, participants were also asked to provide the data logged during their playthrough of the experiment.

B. Questionnaire Structure

On the matter of the questionnaire, first came demographic information. This included standard information on age and gender, which may lead to results regarding specific slices of the population. And a LoC assessment [7], to analyse links between player perception of control and the different levels of randomness.

Regarding the evaluation of the boss encounters, each of the selected items corresponded to a specific dimension of the PX relevant to the hypothesis (see Table III). The items were classified on a 5-point Likert scale. All items were adapted either from the Intrinsic Motivation Inventory [8] (IMI), or Rotter's Scale of LoC [9]. Participants were also asked to compare the encounters and outline any standout aspects.

Dimension	Item	Origin
Enjoyment	The boss fight was enjoyable.	IMI
Competence	I felt competent during this boss fight.	IMI
Effort	The boss fight mattered to me and I put a lot of effort into it.	IMI
Tension	I felt very tense during the boss fight.	IMI
Self-Efficacy (+)	My performance could improve with experience.	Rotter's Scale
Self-Efficacy (-)	I cannot develop the skills to win the boss fight.	Rotter's Scale
Control (+)	During the fight I felt in control.	Rotter's Scale
Control(-)	The boss' behaviour did not allow me to plan ahead.	Rotter's Scale

TABLE III
Boss encounter evaluation items.

C. Preliminary Experiment

This stage focused on assuring that participants of different knowledge levels were capable of carrying out the experiment correctly without guidance or supervision. This design choice was prompted by COVID restrictions and the expediting of data collection. Moreover, it allowed for the balancing of player starting resources in response to the lack of a level preceding the boss fights.

1) Resulting Alterations: From the observations, modifications were made to the form and to the mod itself. For the former, these included simplifying language, clarifying steps of the installation and procedure, and adding explanatory images. For the latter, the main issues to resolve were:

- Player starting resources, as they would not acquire items or consumables before facing the boss variants;
- Boss room layout, specifically as it pertained to finding cover without outright separating player and boss;
- Ease of replayability, to allow for repeated experience with a variant, to form more cohesive opinions;
- Decreasing variability of boss actions to simplify behavior under the limited timeframe of the tests.

These, along with several minor implementation bugs, were addressed and a second version was created.

A second batch of tests was run to test all these changes. With the positive feedback from these, and no further significant alterations, we proceeded to the main experiment stage.

D. Main Experiment

The goal of the main experiment of this work was to find evidence that the devised model for the balancing of randomness, within the selected case study, positively impacted the PX. As such, this experiment was tailored towards finding evidence of increased enjoyment, perception of agency, or other such benefits to the player.

Participants were ostensibly recruited through online platforms, particularly where players of the testbed game could be reached. This was complemented with the gathering of some volunteers through personal connections, with sample diversity in mind.

V. Results

A. Demographic Data

In total, data was collected on the playthroughs of 20 volunteers. For each of these participants, their experiment version was randomly assigned upon starting. The demographic data collected was entirely self-reported.

The population can be said to be primarily males in the 16-24 age range. This matches expectations given the testbed game with which this work was conducted. In total there were 16 male, 3 female and 1 participant who prefered not to identify. Regarding age, there was a 21.15 mean (Std. Dev.=6.769, N=20). As players, most volunteers reported that they made time in their schedules to play video games (60%) and were familiar with the testbed game (70%), having played it multiple times.

To conclude, participant's LoC scores were diverse, with 8 individuals leaning towards an external LoC, a further 8 leaning internal, and the remaining 4 showing no definitive inclination. From the 4 5-level Likert scale items, the scores were summed in relation to the External-Internal locus of control scale (i.e.: a negatively worded item related to an internal LoC, will affect the score towards external the more strongly the participant agreed with it), resulting in a LoC score. For these scores (Mean=10.25, Std. Dev.=2.049, N=20), it can be confidently said it is normally distributed (Shapiro-Wilk Sig.=0.384).

B. Playthrough Data Analysis

In this subsection, the variables related to each participant's playthrough will be analyzed. Amongst these, there are self-reported measures and data logged during play. First, descriptive statistics will be included to contextualize information ahead. Then, an analysis will be made on data validity. Finally, analysis that resulted in statistically relevant observations is presented.

From the 20 participants gathered, 5 played sequence 0, 10 sequence 1, and 5 sequence 2. Three participants did not share the data logged by the game correctly, which means analyses on those variables include data from 17 participants only. For example, regarding the tutorial, the data shows 8 individuals played it, 9 skipped it, and for the remaining 3, the data is missing.

As shown, the distribution of players between the three sequences is far from equal. However, when testing for biases resulting from the sequence played, no results supported the notion. Further, grouping the encounters by order of play, two results of note were found. For boss health, a statistically significant difference was found (Friedman $p = 0.028$), particularly for the first encounter in the sequence. Respectively, the average was 116, 57, and 45 health points from first to last (Wilcoxon $p = 0.028$ and $p = 0.046$), which indicates a possible learning effect. Such an effect would support the implementation of multiple sequences. In the self-reported measures there were no significant disparities, which could point to a player expectation of that learning curve. For the PX, only one dimension had a close-to-significant difference. Enjoyment averaged scores of 4.8, 5.35, and 5.30 for the sequence of fights. These results correspond to a Friedman $p = 0.065$, and Wilcoxon $p = 0.056$ for the first and second encounters.

1) Model Analysis Results: From analyzing the collected data on the boss encounters, self-reported and logged, several interesting results surfaced. Within this section [S], [D], and [C], indicate the result corresponds to the sequential, dynamic, and chaotic variant of the bosses.

First, on the impact of balancing randomness for the PX, results revealed there was a statistically relevant difference in perceived competence based on the boss behavior variant, $X^2(2) = 7.042$, $p = 0.030$. Comparing the variants, the chaotic and sequential behavior patterns were the most disparate, $Z = -1.956$, $p = 0.050$. This seems to support that the balancing of randomness affects player perception of competence, and matches the expectation that the chaotic version is the hardest.

Though no other result was statistically significant, one came close to it, as Negative Self-Efficacy, on a Friedman test (see Table IV). Investigating further through a Wilcoxon Signed-Rank test, a statistically significant difference between the chaotic and dynamic behaviors was found (see Table V). As this disparity was observed between the unpredictable and adjusting variants, the result not only supports the notion that unpredictability increases difficulty, but also that, conversely the adapting behavior seemed to succeed in adjusting to the player.

N	20
Chi-Square	5.059
df	2
Asymp. Sig.	.080

TABLE IV
NEGATIVE SELF-EFFICACY - TEST STATISTICS

	[S] [D]	[D] [C]	[C] [S]
Z	-1.134	-2.070	-1.069
Asymp. Sig. (2-tailed)	.257	.038	.285

TABLE V
NEGATIVE SELF-EFFICACY - WILCOXON SIGNED RANK TEST STATISTICS

Beyond this, tests conducted on participants' LoC scores showed multiple correlations with dimensions of the encounters' evaluations. First, LoC scores correlated strongly and positively with the Enjoyment of the sequential variant (Spearman $rho = .593$, $p = .006$). Suggesting that the more internal a player's LoC is, the more enjoyment they derived from experiences with little to no randomness. From the log data, three correlations were found between LoC and dimensions of the playthrough: a positive correlation with player death count (Spearman $rho = .558$, $p = .020$); and, for the chaotic variant, a negative correlation with player health (Spearman $rho = -.513$, $p = .035$) and a positive one with boss health (Spearman $rho = .505$, $p = .039$). The former two seem to indicate a worse performance on the more randomized fight for participants who believe themselves more in control.

	[S] Enjoyment	Death Count	[C] Player Health	[C] Boss Health
Pearson Correlation	.550	.457	-.513	.443
Sig. (2-tailed)	.012	.065	.035	.075
N	20	17	17	17

TABLE VI
LoC Correlations

C. Discussion

Regarding observations on the boss behavior variants, the lower Perceived Competence and higher Negative Self-Efficacy scores given by players to the chaotic variant, both support the notion that unpredictability increases difficulty. Further, the dynamic variant having the lowest Negative Self-Efficacy scores suggests players felt more capable of improving with that variant.

Moving onto correlations with participants' LoC scores, finding a strong and positive correlation between this score and the scoring of Enjoyment for the sequential variant was unexpected. Given that the closer to 20 a score is, the more internal the individual's LoC is, this appears to show that people who believe themselves in control within their lives, preferred to play an experience which reinforced that perception.

Furthermore, there were negative and positive correlations with the player health and boss health for the chaotic variants, respectively. Both of these results could indicate that player's who believe themselves in control, perform worse in circumstances in which those notions are challenged. Without, however, reporting less enjoyment of them. This point would need to be the focus of future research.

Lastly, there was a positive correlation with Death Count. Given this was an overall value, and not tied with a singular variant, it could mean individuals with more internal LoCs are inclined to make more attempts, before moving on.

To conclude, this work did not find conclusive evidence supporting our hypothesis, but none refuting it either. We also found evidence of players with disparate Locus of Control scores approaching and experiencing the game differently, which should be taken into account in future research. As such, it would be particularly important to re-examine the possibility with both a larger sample size and, perhaps, an approach using starker behavior variants.

VI. Conclusions

This work presents a possible approach to the dynamic adjustment of randomness within a game. Different players have varied play styles, and after extensive hours dedicated to one game, a player can shift play styles, making balancing difficult. In an attempt to manipulate the magnitude of randomness the player faces, finding an adjusting measure was necessary. To that end, we looked at performance and resolved to estimate it through player and adversary resources (in this work, health points). This approach was formulated after an analysis of randomness and its application in game design, as well as the most influenced aspects of the PX and related player traits (LoC).

In this work's experiment, participants faced the three boss variants in one of three different sequences each. These altered boss behaviors include dynamically adjusted behavior, an entirely unpredictable variant, and a deterministic alternative. The latter two were included based on the experiment design, aiming to provide contrast for participants through two static approaches. At the start of a playthrough, one order was randomly assigned to avoid introducing biases in the data. And, after fighting each boss as many times as they chose, players reported on their experience through items adapted from IMI and Rotter's Scale.

By analysing the collected data on the self-reported measures, a positive connection between 'Perceived Competence' and playing the predictable boss variant was found, supporting the expectation that predictability of actions is connected with difficulty. Moreover, the dynamic variant's adjustment seemed to make players feel more capable of improving their skills. As for participant's LoC, an internal mindset was connected with higher enjoyment of the predictable version, worse performance when facing the chaotic boss, and a higher number of attempts at the confrontations.

Overall, the results obtained showed promise in the approach outlined. With that, it is important to consider that as there were a total of 20 participants, in regards to LoC-related observations, further investigation with more participants is warranted. And, on the design of the dynamic variant, feedback from participants in comparing the variants indicates that it may change too subtly, especially between its 'moderate' and chaotic profiles.

A. Future Work

As previously mentioned, in the execution of this research, multiple threads arose which elicit further investigation. These include:

- broader testing of potential links between player LoC and player performance in experiences of varying design, specifically concerning player control (i.e.: level of randomness, freedom of choice);
- investigation of the impact of a starker or subtler profile change, in regards to RNG magnitude, could have on the reported experience;
- exploration of the extension of the model to multiplayer scenarios;
- testing of the application of this or analog models within other genres of game;
- research the impact of the conceptual model within other mechanics, such as reward attribution;

REFERENCES

[1] M. Bar-Hillel and W. A. Wagenaar, "The perception of randomness," *Adv. Appl. Math.*, vol. 12, no. 4, p. 428–454, Dec. 1991. [Online]. Available: https://doi.org/10.1016/0196-8858(91)90029-I

[2] R. Nickerson, "The production and perception of randomness," *Psychological Review*, vol. 109, no. 2, p. 330–357, May 2002. [Online]. Available: https://doi.org/10.1037/0033-295X.109.2.330

[3] J. Schell, *The Art of Game Design: A Book of Lenses*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 2008.

[4] C. Klimmt and T. Hartmann, "Effectance, self-efficacy, and the motivation to play video games," *Playing Video Games: Motives, Responses, and Consequences*, pp. 133–145, 01 2006.

[5] A. Denisova and P. Cairns, "Adaptation in digital games: The effect of challenge adjustment on player performance and experience," in *Proceedings of the 2015 Annual Symposium on Computer-Human Interaction in Play*, ser. CHI PLAY '15. New York, NY, USA: Association for Computing Machinery, 2015, p. 97–101. [Online]. Available: https://doi.org/10.1145/2793107.2793141

[6] R. Hunicke, "The case for dynamic difficulty adjustment in games," in *Proceedings of the 2005 ACM SIGCHI International Conference on Advances in Computer Entertainment Technology*, ser. ACE '05. New York, NY, USA: Association for Computing Machinery, 2005, p. 429–433. [Online]. Available: https://doi.org/10.1145/1178477.1178573

[7] A. Kovaleva, *The IE-4: Construction and Validation of a Short Scale for the Assessment of Locus of Control*, ser. GESIS-Schriftenreihe. Köln: GESIS - Leibniz-Institut für Sozialwissenschaften, 2012, vol. 9.

[8] E. McAuley, T. Duncan, and V. V. Tammen, "Psychometric properties of the intrinsic motivation inventory in a competitive sport setting: A confirmatory factor analysis," *Research Quarterly for Exercise and Sport*, vol. 60, no. 1, pp. 48–58, 1989, pMID: 2489825. [Online]. Available: https://doi.org/10.1080/02701367.1989.10607413

[9] J. B. Rotter, "Generalized expectancies for internal versus external control of reinforcement," *Psychological Monographs: General and Applied*, vol. 80, no. 1, pp. 1–28, 1966. [Online]. Available: https://doi.org/10.1037/h0092976