Are You Lucky or Skilled? An Analysis of Elements of Randomness in Slay the Spire

Johan Andersson*, Mikolaj Trojanowski†, Alberto Alvarez‡
Game Lab, Department of Computer Science and Media Technology
Malmö University, Sweden
Email: *johanandersson@outlook.com, †mikolajtrojanowski@ymail.com, ‡alberto.alvarez@mau.se

***Abstract*—Elements of randomness are a very common factor in modern digital games, from simple rolls of a die to complex AI systems. These elements have an impact on how the player experiences a game. We believe that exploring the field of luck analysis can benefit designers through a developed understanding of how such elements affect players. In this study, we explore how elements of randomness affect players in the roguelike deckbuilding game, Slay the Spire using data clustering. Three player skill groups were identified with the use of clustering: Winners, Low skill losers and High skill losers. Our results indicate that people who succeeded in beating the game, had an increased amount of randomness in the form of cards by a factor of 1.82. Showing that more skilled players do not shy away from randomness but instead embrace it more than lower skilled players.**

***Index Terms*—Game analysis, Game design, Elements of Randomness**

I. Introduction

Nowadays, Elements of Randomness (EoR) are a common aspect of digital games, from as simple as rolling a die to more complex behaviours of AI systems. These elements are outside of the player's control, to some extent, but can have a great impact on the player experience and the outcome of a game [1]. They tend to be difficult for the designers to balance [2] and their inclusion is a common topic of discussion across game communities often tied up to the subject of players' skill. By knowing in advance how randomness affects certain game mechanics and the player experience could enable player-based feedback for designers and make for better games [3].

While EoR:s exist in games with more or less incidence, these elements are of interest in genres where losing or dying is part of the main mechanics since many of them have high penalty for this. For instance, Roguelike games such as Hades [4] or Slay the Spire (StS) [5], feature this mechanic, which can make including, balancing, or managing existing EoRs a complex endeavour for designers, impacting player's experience [6]. Thus, given how important it is for designers to create fair and engaging games when they feature EoRs [7], further research into this is something that would benefit designers and the game industry as a whole. It would allow for further understanding of what luck is in a game like StS, how it affects gameplay in digital card games and how player skill interacts with luck.

This work examines luck and skill, and aims at identifying the margin between them in games, using Slay the Spire as case study. In StS, the player constructs a deck of cards as they play and use them to battle enemy AI in turn-based encounters. The game features numerous EoR:s such as card draw, which cards the player gets to choose from when building their deck, or the card effects themselves. Games, usually, aim at balancing the use of luck and skills to craft competitive and fair environments [2], [8]. Thus, we expect that cards and relics in StS containing EoR:s will be noticeably present across all skill groups, but as the player skill increases (e.g., players play high ascension levels), their impact should become insignificant. Our research questions are; **RQ1:** Is it possible to identify different player skill groups using clustering? and **RQ2:** To what degree do cards and relics containing elements of randomness affect the outcome of a game across different player skill groups in Slay the Spire?

We use a subset of the data released by StS developers, Megacritic, which contains 77 million detailed runs of the game [9]. A "run" in StS involves everything the player did from the start to either when they finished the game or failed, forcing them to start over. Using this data and clustering, we first identify homogeneous groups of variables featuring players of the same skill. Then, we analyze and explore these clusters to observe the proportion of a player's toolkit when playing the game is comprised of EoR:s across different skill groups, which are also identified through clustering. Correlation analysis is then used to conclude how much these variables affect the outcome of a game, and we identify that high skill players that beat the game, had an increased EoR:s when analyzing their cards.

II. Related Research

Gilbert and Wells [1] discuss that there are two definitions of luck. They discuss luck first as "extra-agential" (outside of the player), where the player outcome varies due to forces out of the players' controls. However, this does not take into consideration the effect of the player decisions, thus limiting the scope of what can be classified as random. The second, broader, definition, following three heuristics. These can be then used to measure how an activity or in this case an EoR is placed on the luck-skill spectrum. Maoboussin [10] discussed that in activities where a high level of luck is present, the following is true:

- The outcome is highly unpredictable.
- Great advantages cannot be achieved through learned, repeatable behavior.
- The amount of ”reversion to the mean” in performance is high.

Luck is then categorized into four different types. **Type 1 luck:** Arising from randomizers such as dice, card, and spinners. **Type 2 luck:** Arising from simultaneous decision making, such as in the Sushi Go game [11] **Type 3 luck:** from human performance fluctuating unpredictably in complex circumstances. **Type 4 luck:** Arising from matchmaking. This definition of luck allows for the incorporation of agential factors. This is important as in real games outside the realm of abstraction, they are always present. It allows us to analyze how different player decisions from different player groups react to luck [10].

The specific EoR:s in games can be divided into two types. The types being input and output. Input randomness is defined as randomness that happens before the player can perform any action. Examples of this are randomly generated maps or the hand that is dealt in a card game. Output randomness is defined as events that happen after the player has completed an action. Examples of this would be the random chance to hit in the video game XCOM, a dice roll deciding an outcome, and the random effect of a played card in a card game [12].

A. The impact of luck and randomness in video games

The positive effects of input randomness involve giving the player variance in scenarios and forcing them to choose how to deal with it, while the negative effects are the opposite where the variance allows the system to present scenarios that are not fair nor favourable to the player. An example of a negative effect would be a generated map that has enemy encounters that would be impossible for the player to beat.

Examples of positive effects of output randomness are; it gives the actions of a player variance, simulates mistakes and challenges the player to think about risk management. These aspects add depth to the player experience by increasing the outcome space. An example of a negative effect would be the player experiencing a lack of control over the outcome. This because it occurs after the player has done their action and is unfair if the effects are severe enough to affect the game outcome. One way to tune the fairness of negative output randomness to the player but still maintain the high stakes perception of their outcome is to make the player think negative outcomes can occur. This would lead to their effects being severe, but actually only positive outcomes would be able to occur [12].

In Slay the Spire the developers found that they improved the enjoyment of the game when they moved almost all the randomness in their game from Output to only Input. They did this because they came to the conclusion that because the effects of output randomness is out of the player’s control, it gave a negative experience for the player [13]. One of the developers’ attempts at battling the negative experience is the introduction of a system where the player gets to pick a starting bonus. This bonus is only presented if they made it to the first boss on the previous run. This encourages the player to play with the hand they are dealt rather than trying to restart to get their best possible random starting preconditions.

Furthermore, when randomness is part of main mechanics in games such as movement, their impact can have bigger consequences and exploring their impact in games and the user experience, and alternatives to them is important. De Mesentier Silva et al. [14] analyze how different decks used for “drawing without replacement” and different dice can be used as mechanics based on Salience, Fairness, Disparity and Obfuscation, with recommendations for designers. Similarly, Isaksen et al. [15] study the effect of dice values measuring win bias, tie percentage, and closeness, focusing on how rule changes can have an impact in balance, which is a core topic discussed by game designers [16]. Yin and Xiao [17] focuses on the random rewards given to players for completing tasks, or in-game purchases in the form of Loot Boxes. Their work reveals player perceptions on these depending on how the reward is provided, its effect, and the way it is acquired, providing suggestions for game designers.

B. Skill in video games

Skill level in a game is a form of assessment within the game to measure how skillful a player is at a game or a task within one. Dawson, in her book [18], describes that the analysis of variables such as skill level and variables surrounding that leads to better understanding of the player. To reveal the learning process of a player and what variables contribute to the later achieved skill level can be captured using game telemetry such as log files. In order to properly assess a players’ skill level of a certain task, Dawson points out that data describing the learning process and outcome of that process is required. In other words, what skill level is the player at and how did they reach that goal. Mechanics in the game that have the purposes of measuring that are called “assessment mechanics” [18].

Skill is also linked to the concept of flow [19], particularly in games, where the goal is to have players balancing skill and boredom, so they are fully involved and enjoying the process [20]. Luck can be detrimental for the flow state as this could make games feel less skill and more based on randomness, which is part of what we are interested on investigating in this paper.

In the game Slay the Spire, a part of the assessment mechanic would be beating an ascension level as an example. It defines the expected skill level of a player in form of an outcome. One could argue that since the ascension system is based on levels, we can measure the path the player took to a level. By simply assessing what level of ascension they are at. With this variable selection, proper assessment of the skill level of a player within Slay the Spire can be done.

III. Slay the Spire

Slay the Spire is a roguelike deckbuilding digital card game where the player is expected to fail and start over

TABLE I: Ironclad and Colorless cards containing elements of randomness.

Name	Class	Type	Rarity	Cost	Score	Score Upgraded	Effect
Discovery	Colorless	Skill	Uncommon	1	N/A	N/A	Choose 1 of 3 random cards to add into your hand. It costs 0 this turn. Exhaust.
Jack of All Trades	Colorless	Skill	Uncommon	0	N/A	N/A	Add 1 random Colorless card into your hand. Exhaust.
Chrysalis	Colorless	Skill	Rare	2	N/A	N/A	Shuffle 3 (5) random Skills into your draw pile. They cost 0 this combat. Exhaust.
Magnetism	Colorless	Power	Rare	2 (1)	N/A	N/A	At the start of your turn, add a random Colorless card into your hand.
Metamorphosis	Colorless	Skill	Rare	2	N/A	N/A	Shuffle 3 (5) random Attacks into your draw pile. They cost 0 this combat. Exhaust.
Transmutation	Colorless	Skill	Rare	X	N/A	N/A	Add X random Colorless cards into your hand. They cost 0 this turn. Exhaust.
Violence	Colorless	Skill	Rare	0	N/A	N/A	Put 3 (4) random Attacks from your draw pile into your hand. Exhaust.
Sword Boomerang	Ironclad	Attack	Common	1	98	100	Deal 3 damage to a random enemy 3 (4) times.
True Grit	Ironclad	Skill	Common	1	117	92	Gain 7 Block. Exhaust 1 card at random.
Infernal Blade	Ironclad	Skill	Uncommon	1	192	0	Add a random Attack into your hand. It costs 0 this turn. Exhaust.
Juggernaut	Ironclad	Power	Rare	2	138	26	Whenever you gain Block, deal 5 (7) damage to a random enemy.

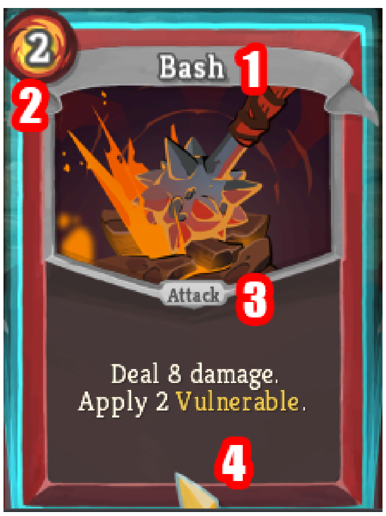


Fig. 1: Representing the terminology of cards in Slay the Spire. 1. Card name. 2. Card cost. 3. Card type. 4. Card description.

numerous times while improving their skill and knowledge of the game, while unlocking new content consisting of new cards, characters, and relics. The goal of the game is to traverse a map of encounters and win those encounters by the use of cards and different powers. The game features four different characters, each having a unique set of cards tied to them. In total the game features 200+ implemented cards, 50+ unique combat encounters and 100+ different items. Each run of the game is very different from each other since the map is procedurally generated and the cards, encounters, and relics vary.

We chose Slay the Spire because of its unique combination between roguelike, deckbuilding, and turn-based gameplay, adding different type of decision-making, its popularity - around 10k active players per day, and the unique opportunity of using the public dataset collected and released by Mega Crit Games [9].

Our analysis focuses on **the Ironclad** and its set of cards and relics. This character class was chosen because it is the first character players have access to and does not require unlocking, which means that most runs will feature it and players of every skill will have experience with it. The set together with colorless cards contains 124 cards. 11 of these cards contain EoR:s (listed in table I), which means that the deck has around 8.9% EoR cards.

Furthermore, *ascension* is a game mode within Slay the Spire that has the purpose of increasing the difficulty of the game by applying different modifiers to the game. The game mode is divided into different levels ranging from 1 to 20, and each level is cumulative, meaning that previous modifiers are applied as you increase level. This is a useful way of measuring the skill level of a player because a higher ascension level reflects a higher difficulty level of the game and shows what experience the player has with the game.

A. Elements of randomness in Slay the Spire

It's important to note that within a game like Slay the Spire, all EoRs where the source is not a card or to a degree relics will be of type input since the player only interacts with the mechanics of the game with the use of cards and nothing else. This results in that the player can often control what EoR:s are in play at a certain time. The player has several tools at their disposal that can manipulate EoRs either by minimizing their effect or completely negating them, for instance, relics that negate these effects or removing cards from the deck.

Although Slay The Spire contains all the types of luck, the focus of this study will be in researching the effects of type 1 luck. The remaining types will be brought back for discussion and future work. **Type 1 luck** refers to cards and relics. Cards and Relics are integral parts of Slay the Spire and provide the massive variability of the gameplay. Their behaviour differs noticeable from each other. Table I contains descriptions of all the Ironclad specific cards and Colorless cards which every character can use. **Type 2 luck** could refer to the process of selecting which cards to play and which to throw into the discard pile. **Type 3 luck** could be exemplified by trying to pull off a complex card combination in order to survive. Finally, **type 4 luck** can refer to encounter randomly selected bosses. These bossed vary in difficulty, hence one can be considered to be ”lucky” to encounter an easier boss.

IV. EXPERIMENT

We only focus on EoR:s found in card and in relics within StS (type 1 luck) because within a card game like StS, cards are the main method with how the player interacts with the game. Analyzing cards and relics allows us to relate more to how the average player interacts with the game and we believe is sufficient enough to get us some measurable results. Relics are also selected for the analysis as they have an impact on how the game is played and affects the players decision making when playing. They are also mainly obtained through the means of input randomness which is another dimension for our study of luck.

We extract a sample from the dataset of 77 million players provided by the developers [9] that we then use to analyze EoR:s. We use clustering algorithms to prepare the data for the analysis and sort players of similar skill into different groups or clusters. The level of skill is determined by looking at similar characteristics between members of the group and comparing them to the characteristics of other groups. These characteristics include: highest level of ascension reached, total score, if the run resulted in a win, hp per floor throughout the run and player card deck score.

The data is processed by dimension reduction algorithms: PCA, T-SNE, and ICA, and we test different clustering algorithms: K-Means, DBSCAN, Spectral Clustering and Agglomerative Clustering. This is necessary because different clustering algorithms are appropriate for different types of data and their use affect how easy it is to conduct observations and draw conclusions from the data [18]. The resulting clusters of players are used for the analysis of the data where players of different skills are compared by how they interact with EoR:s.

The data within these clusters is then analyzed through correlation analysis which consists of measurements between two variables and analyzing the relationship between them. These variables either are dependent or independent. Dependent variables are variables that have an effect on the value of one or multiple other variables. Independent variables do not have any known relationships to other variables. In the case of this study, the dependent variables are variables that measures proportion of EoR in the player’s deck and among relics. The independent variables being analyzed are those that indicate the skill of a player. These include how far the player got in the run, the outcome of the run and how well they are doing at the time. However, it’s important to note and acknowledge that analysis of isolated variables does not reflect how the system behaves as a whole.

V. RESULTS

A. Data

The categorization of the EoR:s into Input and Output randomness is done by identifying what type of randomness the variable is, these are explained under section III and their resulting assignment can be seen in table II.

TABLE II: General elements of randomness in Slay the Spire.

Description	**Variable**	**Type of EoR**
The cards both with and without random effects.	master_deck	Output
The relics both with and without random effects.	relics_obtained	Output
The relic awarded from a battle won against an elite enemy.	relics_obtained & path_taken[E]	Input
The awarded relic when opening a chest.	relics_obtained & path_taken[$]	Input
The relic awarded from a random encounter.	relics_obtained & path_taken[R]	Input
The three available relics to pick from after defeating a boss.	boss_relics	Input
The selected random starting bonus after having defeated the first boss on the previous run.	neow_bonus	Input
The path to the boss with the random event at question mark nodes hidden.	path_taken	Input
The type of encounter found at the question mark nodes on the path to the boss	path_taken[?] & path_per_floor	Input
The random events and the player’s choice with reward.	event_choices & path_per_floor[?]	Input
The three pickable random cards presented when a battle is won.	card_choices	Input
The order of the draw pile which the player draws from at the beginning of every turn.	N/A	Input

We use a sample of *16 638* runs from the 77 million runs. We restricted the data to ascension runs from November 2020, played on the game build ”2020-07-30” with the Ironclad character. The ascesion level distribution had a majority of the runs at level 20 (close to 2700 runs), followed by level 1 (close to 2000 runs), and level 2 (close to 1000 runs). The rest of ascension levels were all between 500 and 1000 runs.

B. Data selection

To determine which variables are selected from the data, a manual inspection of the variables is done to determine which ones are relevant to support our analysis. We selected the variables master_deck and relics_obtained, since they relate directly to the EoR present in cards an relics. To better determine skill we selected: current_hp_per_floor, floor_reached, score, victory and ascension_level. Some data preparation was necessary for clustering to be possible and the methods use to perform this may have also had impact on the result. The *current hp per floor* variable was condensed into a single average number for the whole run because different runs had different lengths and to ensure that all runs were assessed on similar terms. While the average of a series of values is a good way to get a general idea of a run it does not provide the full detailed picture. The same questions can be raised about the *deck score* variable which used the master deck list of cards, translated them into a score and calculated the average for the very same reason as the *current hp per floor variable* where different runs had different lengths of cards.

C. Cluster algorithm

The first step consists on identifying clusters in the data by using different cluster algorithms. Our aim is to cluster based on skill such as that we can then analyze the EoR:s that are present in these. In total, 305 different clustering setups were run and analyzed based on their internal indices and qualitative inspection. It involves a comparison of 4 different clustering algorithms: Agglomerative Clusters, DBSCAN Clusters, K-Means and Spectral Clusters. Evaluation of these clustering algorithms was done by comparing the Calinski-Harabasz Index, Davies-Boulder index and Silhouette score. From the best performing algorithms, Agglomerative and DBSCAN clustered the data into two clusters, winners and losers, but missed the skill nuance. K-Means (K=14) gave too much granularity for both winners and losers, which was unnecesary for this study, although interesting for future work. K-Means (K=5), similar to K=14, gave an extra granularity within the loser side, with a side divided into low and high skilled losers that died quickly, another side similarly divided but the players died late in the game, and a cluster of winners. **K-Means (K=3)**, which we selected, divided the space into three clusters, winners (cluster 1), low-skilled losers (cluster 0) and high-skilled losers (cluster 2), which gave us the necessary perspective and granularity to analyze the EoR:s. The clusters and avg. values within the cluster can be seen in figure 2.

D. Correlation analysis

Having selected the most appropriate clustering algorithm for the purpose of this part of the experiment, the resulting clusters are used to perform correlation analysis. The selected variables are categorized into *dependent* and *independent*. *eor_card_proportion* is the proportion of cards in the player's deck that feature EoR:s and *eor_relic_proportion* is the proportion of relics in the player's inventory that feature EoR:s, and they are both *Dependent*. Hp per floor, floor reached, score, victory, master deck, and ascension level are all *Independent* variables. The variables are analyzed using the Pearson Product Moment Correlation (PPMC). For both the the proportion of cards in the player's deck that feature EoR:s and the proportion of relics in the player's inventory that feature EoR:s there was no strong correlation with each cluster.

While no correlations were strong, there are still notable differences between the different clusters when inspecting the proportion of cards in the players deck. Cluster 1, the group which won the run, had close to zero correlations for all variables except for the master deck score, while the other two clusters had stronger correlations for all variables except ascension level. The extracted correlations between the proportion of relics featuring EoR:s in a run and the player skill variables were similar to the card correlations. No strong correlations were observed when using the proportion of relics, and the differences between the clusters were mostly the same ratio. One notable difference to the card correlations was that relics had lower correlations to player skill variables than which cards had to player skill variables across the board. This could be because the impact of the relics on gameplay is lower than cards and the player has less control over how they are used.

E. Data evaluation

We extracted the specific data that different clusters contain. Specifically, the amount of EoR.s featured in cards and relics within each cluster. We looked at the avg. percentage of (1) cards containing EoR:s in a deck in each cluster, (2) relics containing EoR:s in a run in each cluster, and (3) card containing EoR:s picked when the player had the opportunity to build upon their deck for each cluster. The results are shown in table III.

TABLE III: The proportion of elements with EoR in a run per cluster. Cluster 0 being Low-skill losers, Cluster 1 being Winners and Cluster 2 High-skill losers. The first two rows are proportion of cards and relics with EoR, the next two rows are the proportion of Ironclad or Colorless cards with EoRs, and the last row shows the proportion of cards with EoR chosen by players.

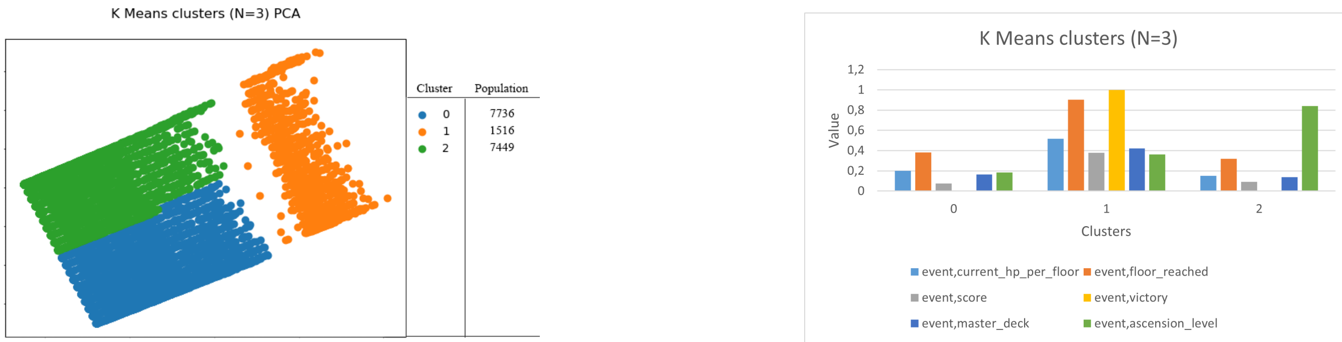
Variable	**All**	**Cluster 0**	**Cluster 1**	**Cluster 2**
Card proportion	2.5%	2.4%	4.3%	2.4%
Relic proportion	6.5%	6.6%	7.7%	6.3%
Ironclad card proportion	2.4%	2.2%	3.9%	2.2%
Colorless card proportion	0.1%	0.1%	0.3%	0.1%
Card proportion	4.9%	4.6%	5%	5.2%

VI. Analysis and Discussion

A. EoR impact

This section describes the impact of EoR:s found in cards and relics on the outcome of a game. The study compares the impact of EoR:s on the different clusters. First describing the general effect and later describing the effect on each cluster. It's important to note that the study looked at the proportion of cards and relics containing EoR:s in the whole deck at the end of a run, but also the proportion of EoR cards chosen by the player as they built their deck.

1) Cards: From the results we can see on average 2.5% of a deck at the end of a run consist of cards containing an EoR. This is a relatively small amount, but still having a bigger difference between clusters than expected. We can attribute the average being low to the fact that the base deck doesn't contain any EoR:s and that a lot of players lose early on, not expanding their deck. Another important observation here



(a) K-Means (K=3) PCA clusters

(b) Avg. values in each cluster.

Fig. 2: Selected clusters for the analysis of EoR:s. (b) shows the avg. values in each of the clusters showing cluster 0 and 2 are composed of low-skilled and high-skilled losers respectively, and cluster 1 are winners.

is that the cluster containing winners had an average EoR proportion of 4.3%. Hence another interesting variable to look at is the proportion of cards containing EoR:s from the cards player use to expand their deck. The value for this landing at an average of 4.9% across all clusters for people who at least added 1 card to their deck. Since cards containing EoR:s make out 8.9% of the whole ironclad set, it's expected that they will make out a smaller percentage of decks as they simply appear less compared to other cards in accordance to the normal distribution. It is why an analysis between clusters themselves is more interesting as we can see the value change under certain conditions instead of just looking at a static value that doesn't allow us to draw a lot of conclusions.

Card Count Looking at the card counts of the different groups in table IV, one can see that as expected the winning runs had a higher total count as those runs lasted longer. What is more interesting is to look at the amount when scaled to the amount of floors reached in the run. In this case it becomes apparent that Winners prefer using smaller decks which makes sense because smaller deck sizes in card games increases the chance of drawing the card you need at that moment. In essence, giving more control over the gameplay.

TABLE IV: Average card count broken down by group. Cluster 0 being Low-skill Losers, Cluster 1 being Winners and Cluster 2 High-skill losers.

Variable	All	Cluster 0	Cluster 1	Cluster 2
Total count	19.37	18.87	28.84	17.95
Count by floors reached	1.15	1.16	0.55	1.28

2) Relics: From the results we can see than on average that 6.5% of relics a player has are relics containing EoR:s. This is a relatively small amount but also expected as the majority of relics in the game do not contain EoR:s. 25 out of 149 (16.8%) relics contain EoR:s that are usable by the iron clad character. Because relics are obtained through the means of random effects besides the shop, it's expected that the EoR relic proportion is low. We can also deduce that the proportion is very equal among the different clusters, besides the winners having a small edge. This we conclude can be attributed to the means of obtaining them.

3) Winners: Diving deeper into the proportion of cards and relics with EoR and their possible impact on the Winners group, a couple of observations can be made.

- The proportion of cards with EoR is higher by a factor of 1.82 in the Winners group than either of the other groups while for relics with EoR, as previously mentioned, were similar for all three groups. This difference between cards and relics with EoR could be down to the fact that the player is offered more opportunities to get new cards and cards have higher impact on the game than relics.
- The proportion of cards with EoR that were chosen by the player is 5%, which is very equal to the other groups. We conclude this in winners not having a preference in cards when looking from a perspective of randomness.
- Ironclad cards with EoR are more popular than Colorless cards with EoR acrossed all three groups. Which is surprising given that there are 7 Colorless cards with EoR while just 4 Ironclad with EoR as seen in table I. An explantion for this might be that colorless cards are only obtainable in special circumstances like from shops and events, while Ironclad cards are rewarded after every combat encounter.
- Looking at the distribution of ascension levels within the Winners group, ascension level 1 is the most popular by far, with level 2 being second, but level 20 is a close third. Indicating that there is a dedicated portion of the player base that makes it to level 20 and keep playing and winning at that level. These people are the best players of the group.

4) Low- and High-skill losers: Comparing the proportions of cards with EoR and the player chosen cards with EoR between the High-skill and Low-skill losers, one can see that both groups have an equal amount of EoR:s in their deck. When compared to winners, these groups underperform quite a bit in the total proportion of cards. This reluctance of cards with EoR among the High-skill losers group is interesting.

Based on this, it seems that as players get more skilled they go from picking cards with EoR:s, to avoid picking them and then back to picking them when they start winning. This shows some kind of progression in terms of player skill, where they are increasing their understanding of the EoR:s in cards and how can be used to the player's advantage.

B. Input and output randomness and their effects

Concluding the previous sections of the analysis we can clearly see that output randomness, in form of card and relics don't seem to have a very big impact on outcome of game in the grand scheme. As a proportion of 4.9% for picked EoR:s means that people still pick 95.1% non EoR cards when provided with a choice. It is interesting as we cannot deduce the reason for this with a big certainty. Is it because people don't like cards containing EoR:s? Or is it because they don't appear as a choice as often compared to other cards? This is set for future work.

What certainly can be deduced is that cards containing EoR:s are a vast minority in the Ironclad set as mentioned previously. It's although interesting how these proportion values go up by a factor of 1.82 which is almost double when comparing wins to losses. It means that output randomness in form of cards does play some role in the outcome of a battle in the game when clusters are compared. This essentially could mean that the downsides of output randomness are far lesser than the potential gain of that randomness. Therefore, the downside of the output is not punishing enough to make an impact and the potential highest value of the output is far better than other competing cards of the same type. Another case for the higher popularity of the cards could be the flexibility of them. For example a card that is defined as "Hit a random enemy for X amount of damage 3 times" might be considered weak against multiple targets, but very strong against one. Meaning that the scenario for when a card is played matters. This study does not take this into consideration and is a topic of potential future work within the research.

1) Win rate analysis and input randomness: During the clustering process we could draw out from the data the average win rate in the game. The win rate was around 9%. This means that 1 in 10 runs results in win, which could point out that there is either a high level of randomness, a high skill cap or both. Slay the Spire is a rogue-like game, meaning that essentially you are expected to fail in order to get better on your next run. While isolating this win rate, we can certainly say that the developers succeeded in that goal. The question is what measures did the developers take to achieve this? From our analysis, we can observe that there is both a high amount of input randomness involved but also a high skill cap. The game has 10 EoR:s of the input type (for a description of these see table II) in play at all times, and more are introduced as the ascension level goes up. This results in a huge variety between each run and how it plays out. In short, this means that every player in this game has different start points, which they have to overcome in order to win. The start point might be difficult or easy but so can the path to victory. This requires a high adaptability level from the player to uncertain situations, which, if succeeded, is a sign of a skilled player. Based on the study, we believe that input randomness has a big impact on the outcome of a game like Slay the Spire, but it requires a player to make adjustments and make room for elements of output randomness in order to have a higher chance of succeeding.

C. Takeaway for Game Designers

Our analysis shed light into the use and effect of EoR:s, and how diverse set of skilled players harnessed those. We expect that this research can help inform (or at least, put the focus for) game designers about the impact of these EoR:s, both in Slay the Spire and in general, for balancing purposes. Balancing is one of the hardest tasks designers are tasked with, and understanding, firstly, what type of EoR:s exist, and secondly, their impact across set of skilled players is paramount. One goal with this kind of analysis could then be to allow designers to create predictive models for how EoR:s affect player experience and play styles across different skill groups similar to how other games focus on balancing this [8]. Work such as the one by Lindstedt and Gray [21] that focus on understanding expert behavior could be complemented with how different game elements and factors such as EoR:s are utilized by players, which could also be compared to what designers expect from them.

D. Limitations

Our study is just a stepping stone in this area. We only use Slay the Spire to explore this, but our analysis would have to be expanded to more games; preferably to games where the community view randomness as an issue or just being a big part of the game play to get more general results, conclusions, and guidelines. Within Slay the Spire, our study is limited both by the narrow approach taken by analysing only **the Ironclad** class, and by the cluster division we did. Finally, our analysis is based entirely on the game data provided by the developers [9], but this should be complemented with qualitative studies to understand better the player experience and how it is affected by the choice of EoR:s, which is left for future research.

VII. Conclusion

The evaluated clustering algorithms were proved viable for identifying different player skill groups and sorting the players into these groups in Slay the Spire. From these groups, it was also possible to extract the information necessary to analyze EoR:s in cards and relics. In general, input randomness has a big impact on the outcome of a game like Slay the Spire. As it's reflected in both the average win rate of players at 9% but also how varied the data is in each cluster with the main example of this being what encounters players face. The game has a minimum of 10 input variables in play at all times and that number increases with the difficulty level selected in form of the ascension level. It's important to note that all forms of input randomness do not have equal impact. For example we note that relics do not have a strong correlation to the player skill variables. We also conclude that it requires a player to make adjustments and make room for elements of output randomness in form of cards and relics. This in order to have a higher chance of succeeding as we observed the varying level of randomness proportion in groups of winners compared to losers where the former has a proportion 1.82 times greater than the latter.

Players of higher skill handle the elements of randomness featured in cards better than players of low skill. The study shows that players of all skill levels interact with cards and relics featuring elements of randomness and are affected by them. There is a preference among players that win at the game to pick cards featuring elements of randomness, showing that more skilled players know they can overcome the possible negative elements of randomness of the cards and use them to their advantage. Skilled in this regard means that they succeeded at beating the game. With regards to relics featuring elements of randomness there is no discernible difference in preference between players of low and high skill and no measurable difference in how they handle them. This because the player has no effect on how relics are obtained resulting in an even distribution among groups.

VIII. Future Work

We view this study with potential for future work in order to expand the field of luck analysis in video games and understand their effects in Slay the Spire. Our study focused on K=3, dividing the cluster into low-skill losers, high-skill losers, and winners. One expansion could be to use more granularity to gather a deeper insight into the possible groups. This could show how winners compare to each other across different ascension levels Further, this study focused only on one class within the game. Future work where a similar study should analyze different classes and compare the results between them. It's also desirable that a wider range of EoR:s and skill variables can be analyzed in order to draw more general conclusions. This study focused on the "Hard mode" of the game to gauge how skilled a player was. Perhaps there are better ways of measuring how skilled a player is and that in itself is a huge area for future studies to explore, "How do we know a player is skilled?"

We believe that an analysis of how cards synergize with other card would be benefiting to this kind of analysis. Interesting factors to look at is how cards featuring EoR:s get picked based on their synergy score among the different skill groups. Our hypothesis for this is that groups of higher skill would pick cards with a higher synergy score than players of lower skill. This would enable for a quantitative way of measuring card effect on player preferences. Further, it would be interesting to know if these cards are more flexible in general, making them better candidates for multiple deck builds. We focused on type 1 luck, randomizers. It would be beneficial to the analysis to look at the the other 3 types, simultaneous decision making, human performance and matchmaking. Comparing the effects of the 4 types to see which one has the biggest effect on the player within Slay the Spire. This in order to then narrow down the analysis on the type with the biggest effect.

Finally, since the goal of the study aims to provide an understanding for designers on how their elements affect players. It would only be logical to ask the players how they feel about the subject matter. Our study provides numbers to how big the proportion of EoR:s are, but can't explain if that's a lot or not in the context of a player, or in what way do these EoR:s influence the player experince. It would help to conduct interviews with players in regard to how they feel about EoR:s in their game and how it affects them. This would allow us to draw more detailed conclusion and allow us to tell if 9% win rate is too little, or if decks with an EoR proportion greater by a factor of 1.82 are the ones that win is optimal.

References

[1] D. E. Gilbert and M. T. Wells, "Ludometrics: Luck, and how to measure it," *Journal of Quantitative Analysis in Sports*, vol. 15, 2019.
[2] T. Fullerton, *Game Design Workshop: A Playcentric Approach to Creating Innovative Games*. 2008.
[3] B. G. Weber, M. Mateas, and A. H. Jhala, "Using data mining to model player experience," 2011.
[4] Supergiant Games, "Hades," Dec. 2018. Game [PC]. Supergiant Games, California, US. Last played January 2021.
[5] Mega Crit Games, "Slay the Spire," Nov. 2017. Game [PC]. Mega Crit Games, Washington, US. Last played March 2023.
[6] A. Drachen, P. Mirza-Babaei, and L. Nacke, *Games User Research*. Oxford University Press, 2018.
[7] W. IJsselsteijn, Y. D. Kort, K. Poels, A. Jurgelionis, and F. Bellotti, "Characterising and measuring user experiences in digital games," *International Conference on Advances in Computer Entertainment Technology*, vol. 620, 2007.
[8] C. Douglas, *Luck and Skill in Professional League of Legends (E-sports)*. PhD thesis, MIT, 2018.
[9] Mega Crit Games, "Slay the Spire - 77 million run," 2020.
[10] M. J. Mauboussin, "The success equation: Untangling skill and luck in business, sports, and investing," *CFA Institute Conference Proceedings Quarterly*, vol. 30, 2013.
[11] Phil Walker-Harding, "Sushi Go!," 2013. Game [Tabletop]. Phil Walker-Harding, Sydney, AU. Last played January 2021.
[12] "Randomness and game design – keith burgun."
[13] "Why revealing all is the secret of slay the spire's success — rock paper shotgun."
[14] F. de Mesentier Silva, C. Salge, A. Isaksen, J. Togelius, and A. Nealen, "Drawing without Replacement as a Game Mechanic," in *Proceedings of the 13th International Conference on the Foundations of Digital Games*, FDG '18, (New York, NY, USA), Association for Computing Machinery, 2018. event-place: Malmö, Sweden.
[15] A. Isaksen, C. Holmgård, J. Togelius, and A. Nealen, "Characterising Score Distributions in Dice Games," *Game & Puzzle Design*, vol. 2, no. 1, pp. 24–37, 2016.
[16] T. Fullerton, C. Swain, and S. Hoffman, *Game Design Workshop: Designing, Prototyping, & Playtesting Games*. Gama Network, Taylor & Francis, 2004.
[17] M. Yin and R. Xiao, "The Reward for Luck: Understanding the Effect of Random Reward Mechanisms in Video Games on Player Experience," in *CHI Conference on Human Factors in Computing Systems*, (New Orleans LA USA), pp. 1–14, ACM, Apr. 2022.
[18] C. Dawson, *Game analytics*. 2019.
[19] M. Csikszentmihalyi, *Flow: The Psychology of Optimal Experience*. New York, NY: Harper Perennial, Mar. 1991. Published: Paperback.
[20] M. Klarkowski, D. Johnson, P. Wyeth, S. Smith, and C. Phillips, "Operationalising and Measuring Flow in Video Games," in *Proceedings of the Annual Meeting of the Australian Special Interest Group for Computer Human Interaction*, (Parkville VIC Australia), pp. 114–118, ACM, Dec. 2015.
[21] J. K. Lindstedt and W. D. Gray, "Extreme Expertise: Exploring Expert Behavior in Tetris," *Cognitive Science*, vol. 35, 2013.