The Effect of Leaderboard Ranking on Players' Perception of Gaming Fun

Charles Butler

The Norwegian School of Information Technology, Oslo, Norway
charlesabutler@gmail.com

Abstract. Although fun is desirable in nearly all commercial games, defining it and actually getting it into a game can prove difficult. Developers have added multiplayer features to their games since the beginning of the industry in an attempt to create fun, but to what extent does this actually affect a player's perception of a game's fun? This paper gives an overview of relevant research relating to fun and play before attempting to tackle the key issue of the effect of player success as measured by leaderboard rankings on the perception of a game's fun.

Keywords: fun, play, video games, computer games, game design.

1 Introduction

There are a number of difficulties when attempting to study fun or people's perception thereof. It seems problematic to attach a concrete definition to such an abstract concept. Additionally, since most people likely feel very familiar with the concept, finding a definition that all can agree on is even more challenging. Few studies tackle the concept directly, but there is a body of research concerning the concept of play. Since it seems logical to associate the activities of play with the concept of fun, it also seems valid to consider these activities as well. Much of this research involves attempting to classify or categorize various activities to form a model of play.

A common trend in many modern games is the addition of various multiplayer features. These often seem incidental to, or even entirely separate from, the main gameplay of the game, even though they can certainly add replay value and additional interest. A notable example of this is the Xbox Live functionality that integrates with Xbox 360 games. Of course, some games benefit tremendously from this added functionality, but do they all? Designers must be aware that certain features can both add and detract from the perception of fun if they are to make informed decisions about what features to add.

This project investigated the effect of one specific game alteration on the perceived fun of players. The testing process allowed testers to play a game that randomly selected one of three leaderboard versions to present to the player at the game's conclusion. One leaderboard version was non-populated, having only the player's score listed. The game dynamically populated another version with names and scores higher

A.A. Ozok and P. Zaphiris (Eds.): OCSC/HCII 2013, LNCS 8029, pp. 129–136, 2013.

than the player's score, and the last used names and scores lower than the player's score.

It may seem that a more successful player would perceive themselves to be having more fun than a less successful player and vice versa. However, this project set out to test whether or not leaderboard results would alter a player's perception of fun and to what extent. The concept of a leaderboard serves as a proxy for success relative to that of other players and could be of use to game designers in determining which multiplayer features to incorporate into their designs. While many different factors likely influence the feature set of a game, additional data concerning this specific factor could provide extra insight in certain situations, even if it is not the primary motivation.

2 Background

A review of the common background material related to fun and play investigated a number of different areas in an attempt to understand how players derive a sense of fun from games.

Salen and Zimmerman [8] examined a number of categorizations in an attempt to classify either the type of activity that leads to pleasure in play, or fun, or the resulting emotion. One common, and incredibly broad, definition of fun is offered:

> *"Game designer Hal Barwood organizes all of the varied emotions a game can produce under the heading of "fun." Fun games are what players want. A fun game makes for a pleasurable experience, which is why people play."*

However, game designer Marc LeBlanc [8] finds the term insufficient to fully describe the concept. He instead developed a categorization based on of the type of pleasure that players experience while playing games. Psychologist Michael J. Apter [8] offered a very similar categorization in his essay "A Structural-Phenomenology of Play" that focuses on the cognitive arousal that play provides.

Anthropologist Roger Caillois [8, 3] developed one of the more widely used categorizations of the different forms of play:

1. Agôn: competition and competitive struggle
2. Alea: submission to the fortunes of chance
3. Mimicry: role-playing and make believe
4. Ilinx: vertigo and physical sensation

Additionally, Caillois expands on these categories with the concepts of paida and ludus. Paida and ludus add additional depth to each of his four categories by providing a continuum from pure paida to pure ludus upon which the various games or play types fall.

"Paida represents wild, free-form, improvisational play, whereas ludus represents rule-bound, regulated, formalized play"

However, while the type of play may give clues to the player's motivations, it does not specifically address it. Game designers should be interested in the types of play that people choose to participate as well as their reasons for doing so. Game designers Neal and Jana Hallford [8, 5] point out the importance of rewards in keeping the player satisfied, offering a categorization based on the types of rewards offered to players.

While those rewards may describe various ways that a designer can keep players interested by giving rewards to the players' avatars, individual players value various rewards to very differing extents. People play games for many reasons, and it seems logical to assume that they are predominately doing so because playing gives them some sort of reward. It also seems to follow that the game designer could do a far better job providing fun for the player if it was easy to tell what rewards a player values most. While many games are likely self-selecting (for example, racing games are likely to attract the type of player that feels most rewarded by ilinx types of play), the different categorizations of play presented above do not speak directly to what the player might enjoy. Richard Bartle [1] addresses the motivations behind various player actions and the things designers can do to reward them. After conducting an analysis of the responses to a long running debate about what players want from a game, Bartle categorized players into four quadrants, commonly known as the Bartle Player Types, consisting of Achievers, Socializers, Explorers, and Killers. Later, in an effort to further define variances he noticed within his player types, he expanded the graph to include an additional axis based on whether players interacted with the game world in an implicit or explicit manner.

By analyzing which features appeal to which groups, a game designer can more accurately predict how well a game fits with its intended target audience. The designer may also be able to broaden the appeal of the game by intentionally adding features that appeal to players outside of that audience. Bartle's types deal primarily with the motivations of the players, but he considers analysis of their actual behavior to also be worthy of comment. In regard to in-game behavior, he uses a study conducted by the system administrator of *Habitat*, F. Randall Farmer [1], which categorized players based specifically on their behavior. Farmer also noted that players tended to change behaviors over time. Players would start being primarily interested in the game, but as they became more experienced, they would begin to derive more and more of their fun from their interactions with the community. In a similar fashion, Bartle has noted the ways in which players seem to progress from one of his types to another as they become more experienced with the game [1].

Raph Koster [7] indicates that fun is essentially derived from the player's brain attempting to find patterns and succeeding in doing so, meaning that learning is really the mechanism that allows for fun. However, he sees the need to refine what people mean when they talk about having fun. He compares his categorization of the types of fun to Nicole Lazzaro's [7] four clusters of emotions: hard fun, easy fun, altered states, and the people factor. Koster also notes that we should be mindful of the

positive emotions of interpersonal interactions that are often enjoyable even though they may not necessarily meet his definition of fun. While often overlooked by designers, these are consistent with Farmer's [1] notion of players becoming caretakers of other players as they grow. Allowing for the development of these emotions within a game could be crucial in keeping players satisfied and increasing the longevity of a game.

Koster [7] contends that much of what humans perceive as fun stems from activities that aid in survival and that games naturally evolve along with the needs of the players to serve them in the survival of their species. One notion that follows from the idea of games as an evolutionary teacher is that of the optimal arousal level theory [6]. This concept states that higher mammalian brains naturally pursue an equilibrium level of arousal, seeking out exciting activities when bored or relaxing activities when stressed. This could help explain the differences in what various players consider fun.

Howard Bloom [2] reinforces the need for a certain level of arousal. He argues against the trend of reducing stress and competition in our society. The popular belief is that stress causes health problems, but the benefits of actively engaging in competition, such as increasing one's social status, actually imparts significant health benefits such as lower blood pressure and enhanced emotional states.

Many of the previously cited individuals refer to the work of Mihaly Csikszentmihalyi [4]. His concept of flow is very relevant to fun and, as a result, is sought after by game designers who want their players to experience flow-like states (which could arguably be called synonymous to immersion) while playing games.

3 Method

3.1 The Project

This project involved the creation of a simple, web-based game created with Adobe Flash that has three different leaderboard versions. The game is a side-scrolling shooter that tracks the player's score and displays a leaderboard at the end of the game that manufactures the intended results regardless of the player's score. One version of the leaderboard indicated that the player had a higher score than any of the previous players. Another indicated that the player had a lower score than any of the previous players. The final version of the leaderboard displayed only the current player's score with all other slots being empty. Upon an initial play of the game, a leaderboard version was randomly selected. If a player chose to replay the game or returned to the site later, the page loaded the game with the same version of the leaderboard that the player previously encountered, although without the attached surveys or additional screens. However, even though the leaderboard encountered upon replay contained the same data as the initial version, the replay version was functional, and the player was able to place scores on it naturally.

While the scope of the game intended for this project was very limited, this was not detrimental to the data collection. The player only needed to play for a short time for the collection of data, so replayability was not a high priority. The quality of the

game did not seem to dissuade many testers as the quality of both the visuals and gameplay were on par with Flash games commonly found on the internet.

3.2 Recruiting Testers

The testing used posts on online discussion forums to recruit participants. This method allowed the recruitment of a large number of testers with minimal effort, and it was highly scalable, allowing for the easy expansion of future testing sessions. To test the hypothesis, the only required data was the player's game version and the corresponding fun rating. However, additional data allowed for a more thorough analysis. Demographic data allowed the project to test for variances in extent of the effect on fun based on different groupings of people.

For this test, the recruitment posts focused specifically on forums that cater to video game players. This choice in recruiting assumed that individuals who frequent video game forums would be ideal initial testers for a number of reasons. First, they were assumed to be more likely to participate in the test than users of non-gaming forums. Second, they were assumed to be more comfortable with the technology, conventions, and terminology used in the test, improving the survey responses, and finally, they were assumed to be more capable of successfully handling any technical difficulties that might arise, reducing the abandonment rate.

3.3 Analysis

After filtering out invalid log entries (primarily due to abandoned surveys and technical problems encountered by the testers), complete surveys and accompanying gameplay data were collected on 132 unique testers (repeat players were tracked and filtered by IP address). Of these 132 participants, 54 players received the winning leaderboard version, 39 players received the losing version, and 39 players received the blank version.

There were three main areas of focus in the data analysis for this project. The first area was concerned with the relevant gameplay data collected during the testing sessions, including score, replay percentage, and difficulty rating. The second was an examination of the game ranking metrics and the distribution of those responses across the three different game types. This analysis attempted to detect any patterns or correlations between the game type and the ratings that the participants submitted on their surveys. The third segment of the analysis attempted to determine if any of the additional survey data collected about the participants related in any meaningful way to the fun ratings given to the game. This analysis broke the data down based on the different fun ratings given. (Readers should note that this analysis combined the results of the highest and second highest fun ranking due to their only being one response in the highest rank.)

The analysis first checked the data at a high level to see if any patterns or apparent inconsistencies emerged. Contingency tables served to help analyze the relationship between variables, and their use was appropriate here as the collected data contained a number of potentially related variables. A contingency table analysis using the

chi-square statistic attempted (where applicable) to determine if there was a demonstrable relationship between the two nominal scale variables (i.e. the fun rating and the other rating specific to the test).

3.4 Potential Issues

It was impossible to avoid bias entirely, so a few of the potential biases of particular concern appear here. First, biases based on factors such as demographic differences, gameplay preferences, and the locations of the requests for testers were almost certain to occur. The analysis of the final data attempted to mitigate this bias by grouping similar players to identify any serious biases. Additionally, the chosen testers were of a very specific population for the reasons mentioned above, so to get results representative of the general population, much more extensive testing would need to be conducted.

The number of testers could be a problem as well. While 132 testers is a significant number of testers, once broken down into segments, the numbers are considerably less significant, making the chance of skewed results very high for some comparisons.

The intended perception of the losing condition was likely not communicated in many cases. Scoring high enough to even appear on a typical leader board could be seen as a sign of success and not at all like a negative condition. This may serve to make the results more about varying extents of success than winning versus losing.

When questioning players about the difficulty of the game, the intent of the question should be more specifically defined. It is possible that there was confusion as to whether the question addressed the complexity of the game or the difficulty in conquering the gameplay challenges.

4 Key Findings

- Winning, losing, or the lack of competition does affect the player, though not always in the expected way or in the same way for all players.
- The winning condition seemed to polarize the perception of fun, causing clusters of responses at the upper and lower bounds of the responses.
- The losing condition seemed to centralize the perceptions of fun, causing the responses to cluster in the middle of the scale.
- Players encountering competition (a winning or losing condition) replayed the game at least once approximately 50% more often than those with the neutral condition.
- Players who encountered the losing condition replayed the game at least once approximately 100% more often than those with the neutral condition.
- Even though it likely runs counter to the expectations of many game developers, fun ratings do not necessarily correlate with a player's likeliness of replaying a game (though there may be a threshold above which this changes).

- Measuring behavior is more reliable than measuring responses. The responses given to the replaying likeliness question were not strongly correlated to the measured replay data.
- Making a game too easy or too shallow can hurt its fun rating as judged by the player comments.
- Once a player decides that a game is too easy, it might be difficult to change that opinion, or alternatively, the player may be considering a different definition of difficulty.
- Higher scores do not necessarily correlate with more fun.

5 Conclusions

Presented here are a number of models describing various types of play, fun, and related emotions. While no single model is universally accepted, one can still find many useful and relevant points within the various approaches. In the search for fun, one is likely to benefit from viewing the problem from multiple perspectives, examining the ways in which people play, the motivations behind the desire to play, the specific features commonly cited as fun, and the emotions associated with each. All are valuable and provide insight into creating effective game designs.

The intent of this project was to examine a specific game mechanic for its effect on perceived fun. The hypothesis was that success relative to other players has a corresponding effect on the player's perception of fun. This work offers game designers quantitative evidence useful in determining what effect the addition of related multiplayer features might have on their game. The alteration of a high-score screen was appropriate because it is a feature that has very little, if anything, to do with the actual gameplay of a game, and yet, games have had these screens added to them since the early days of the industry. They add a mechanism for encouraging competition by quantifying "bragging rights" to a game. This is likely to appeal to a certain subset of players, but the potential to affect the other players is important to consider as well.

After the collection and analysis of the data, there were a number of findings. The first concerned the main hypothesis that success relative to other players has a corresponding effect on the player's perception of fun. The findings indicate that success relative to other players does affect the player's perception of fun, though not necessarily in a corresponding way (at least for all players). Success seems to affect players' perceptions in different ways, possibly depending on their expectations and any number of other factors. The results seem to indicate that success has a polarizing effect, increasing the number of high and low fun responses. Additionally, it seems that failure has a centering effect, encouraging more medium responses, and even though it appears to reduce high responses, it seems to reduce low responses by a similar amount.

One of the most telling outcomes was the number of players who replayed the game at least one additional time after completing the survey. As expected, a higher percentage of players receiving the losing version replayed the game than of those with the winning or blank versions. In fact, the losing version's replay percentage was

more than double that of the blank version. A post-game survey question asked the players about the likeliness of playing again. The responses to this question indicate that players of the losing version are only slightly more likely to play again than the players of the other two versions (whose responses were practically equal). This supports the idea that what players say and what they actually do is not always the same. Additionally, the replay percentage did not seem directly correlated with the players' fun rating. The replay percentage of players who ranked the game highest was over 3.5 times higher than the overall average, but the lowest fun grouping replayed the game three times more often than the second and twice as often as the third. It is possible that until a player perceives a game as being "fun enough," any replay that occurs is for a reason other than fun, whereas above that threshold, fun may become more relevant.

One of the most significant unexpected results was the number of players who rated the game as too easy. This was surprising for a game that becomes practically impossible after a few short minutes. It seems likely that designing for accessibility past a certain point reduces the fun of your game to a certain number of players. One possible cause is that some players may quickly decide how easy or difficult a game is and stick to that decision no matter how difficult it becomes. Another possibility is that players may equate difficulty to the ease of learning and playing the game, causing players to rate games with simple mechanics as easy regardless of the difficulty of the overcoming the gameplay challenges.

One other unexpected result was that score seemed to have no significant effect on fun rating. It was an assumption that players better at the game would give it a higher fun score, and while players who played similar game types rated the game as more fun, playing similar games apparently was not correlated with achieving high scores.

References

1. Bartle, R.A.: Designing Virtual Worlds. New Rider Publishing, Indianapolis (2003)
2. Bloom, H.: The Lucifer Principle: A Scientific Expedition into the Forces of History. The Atlantic Monthly Press, New York (1995)
3. Caillois, R.: Man, Play, and Games. University of Illinois Press, Champaign (2001)
4. Csikszentmihalyi, M.: Flow: The Psychology of Optimal Experience. HarperCollins, New York (1991)
5. Hallford, Hallford: Swords and Circuitry: A Designer's Guide to Computer Role Playing Games. Prima Publishing, Roseville (2001)
6. Hebb, D.O.: Drives and the C.N.S (Conceptual Nervous System). Psychological Review 62, 243–254 (1955)
7. Koster, R.: A Theory of Fun for Game Design. Paraglyph Press, Inc., Scottsdale (2005)
8. Salen, Zimmerman: Rules of Play: Game Design Fundamentals. The MIT Press, Cambridge (2004)