REWARD PREFERENCE IN VIDEO GAMES

Kung-Chien Dominic Kao

A THESIS

PRESENTED TO THE FACULTY

OF PRINCETON UNIVERSITY

IN CANDIDACY FOR THE DEGREE

OF MASTER OF SCIENCE IN ENGINEERING

RECOMMENDED FOR ACCEPTANCE

BY THE DEPARTMENT OF

COMPUTER SCIENCE

Adviser: Yael Niv

June 2012

ABSTRACT

The world has already spent 5.93 million years playing the video game World of Warcraft. What makes video game rewarding? We suggest that one important factor in a player's enjoyment of a game is the schedule of rewards, specifically, we hypothesize that variable rewards are more interesting and enticing than a fixed reward size. To test this hypothesis we studied several variable reward schedules contrasted with a fixed reward schedule in a custom-built video game. We found that variable reward schedules were preferred to fixed reward schedules with the same mean, except when the variable schedule's range was extended. We further hypothesize that reward schedule preference is a good indicator for game genre preference. The participants favoring some game genres had strong inclinations for a particular reward schedule. Ultimately, our conclusions can be applied to make more effective the now ubiquitous game mechanics that surround us in both the real and virtual world.

ACKNOWLEDGEMENTS

I would like to thank my advisers Yael Niv and Carlos Diuk for their never-ending support, Adam Finkelstein for his readiness to assist me, Reiner Prokein for his indispensable character sprites, John Hopson, Richard Bartle, Julian Togelius, Daphne Bavalier, and Nick Yee all for providing preliminary feedback. Finally, I want to thank Melissa Lawson for lighting the way, the Princeton Computer Science department for supporting me, and NSERC for their generous funding.

CONTENTS

FIGURES

TABLES

"We're witnessing what amounts to no less than a mass exodus to virtual worlds and online game environments"

Edward Castranova

INTRODUCTION

By the age of 21, the average American will have played 10,000 hours of video games [1]. What makes video games fun? We believe one of the key ingredients in a player's enjoyment of a video game is reward schedules. Reward schedules have been studied extensively in the context of instrumental conditioning, frequently using animals such as rats, inside a so-called Skinner box [2]. These boxes contain, at their most basic, a lever and a food dispenser. The rat inside of the box must pull the lever (respond) in order to be rewarded a food pellet (reward). Whether or not the animal receives food pellets for any specific action depends on the reward schedule applied. For instance, in a *fixed ratio n* schedule, a reward is given after every *n*th response. In a *variable ratio* schedule, a reward is given after a variable number of responses. These reward schedules form a part of animal learning theories [3] [4] [5] [6] [7], which in turn, are the source of our later hypothesis that variable reward schedules are preferred.

Little work exists in this area with games and with humans as test subjects. The aforementioned variable ratio schedules have been shown to produce higher rates of responding than fixed ratio schedules. However, variable ratio schedules may not be *preferred* in the context of a video game, as people may be exhausted or frustrated from such high rates of response or from the unpredictability of reward. Therefore, we sought to measure actual player preferences of reward schedules by allowing participants to choose which schedules they want to play. We did this by creating a video game and allowing participants to select between two game worlds differing only in their reward schedule. We then used both choice data and survey answers to determine whether a preference exists for one reward schedule over another.

Despite a lack of research in this area, game-like mechanics are everywhere around us. LinkedIn uses a progress bar to encourage people to fill in their profile with more information. Panera offers a MyPanera rewards card which is swiped when the owner makes a purchase. Occasionally, this will result in a reward such as a free pastry or drink. It is impossible to predict when the rewards are given (even seasoned cashiers who have worked at Panera for many years have reported that it is random); in fact, the reward program operates under a variable schedule. Many stores offer punch cards which act as fixed reward schedules (buy *n* drinks and get your *n+1* drink free). This is an area ripe for research with wide applicability.

Understanding current video game research gives us peripheral context for our work. In one of the earliest works of video game research, players of online MUDs (Multi-User Dungeons, effectively text-based versions of today's online role playing games) were placed into four categories: achievers (want to attain goals), explorers (want to understand), socializers (want to communicate), and killers (want to kill others) [8]. More recently, Yee (2006) studied thousands of players of MMORPGs (massively multiplayer online role playing games) and found that there were three overarching groups and ten subcomponents to these groups: achievement (advancement, mechanics, competition), social (socializing, relationship, teamwork) and immersion (discovery, role-playing, customization, escapism) [9]. While these works attempt to categorize players, we seek instead to find the reward schedule preferences of players in general.

Romero (2008) provides time-tested industry advice for performing research on a game [10]. He suggests (using Halo 2 from Microsoft as one of the examples) surveying participants often during the gameplay experience, using memorable quotes to link findings, and to respect attitudinal data as much as quantitative data (e.g. “I feel this game was clunky”). This advice is mainly geared towards making an existing game better but provides industry perspective on games research.

The preceding demonstrates that reward schedules and video games is a seldom explored topic in academic research. Nonetheless, research that does exist provides frameworks which may help to explain why people enjoy video games (our original, fundamental question) and a context within which to place our work. We will touch on a few more related works which pertain to our results in the Discussion.

METHODS

We created a point and click action-roleplaying game with similar game mechanics to a commercial game, Diablo 2 (Figure 1). Creating our own game afforded us the flexibility in controlling game parameters and data collection. In our game, participants have two ways of controlling their character. With single left-clicks, participants can move their character to a clicked location, or perform a single attack on a clicked monster. By left-clicking and holding down the mouse button, participants can move their character in the direction of the mouse pointer indefinitely, or perform repeated attacks on a clicked monster.

Figure 1: Commercial game Diablo 2 (top half) contrasted with our game (bottom half)

The game screen which participants see consists of the following key elements (see Figure 2): a game timer (games last one minute each), a large point total (points are gained from killing monsters), a red semi-circle denoting the participant's health (health is lost from fighting monsters), and an experience bar (a visual representation of points earned). After filling the experience bar, participants receive an upgraded character which has both increased attack power and health. There are a maximum of four possible upgrades (Figure 3).

Figure 2: Screenshot from our game showing: Timer (top-left), Point Total (top),
Health (bottom), Experience Bar (bottom)


Figure 3: Character upgrades in increasing power from left to right

The game world (Figure 4) consists of hard obstacles (which cannot be walked on) such as cliffs, rocks, trees and a center fountain. The positions of these obstacles remain fixed from game to game. The fountain in the center of the game world serves as a safe haven for participants where they recover lost health (monsters will run away if they chase the participant close to the fountain). Monsters, on the other hand, change positions from game to game. We created several distributions of monsters, each with similar spacing across the game world, and one of these distributions was chosen at random at the start of each game.

Figure 4: Empty game world (trees and monsters are generated at run-time)

In order to compare reward schedules, we created two worlds, identical in every respect except color and reward schedule (Figure 5).

Figure 5: Choosing between game worlds (each with a different reward schedule)

The experiment can be broken down at a high level as follows (Figure 6): For each participant a fixed and variable schedule were each randomly assigned to either the blue or the green world. The participant then underwent two training sessions, one in each of the worlds in a randomized order. Training consisted of killing a set number of monsters in an enclosed area of the world.

Rewards were set such that participants received the same exact total number of points in the two training worlds.

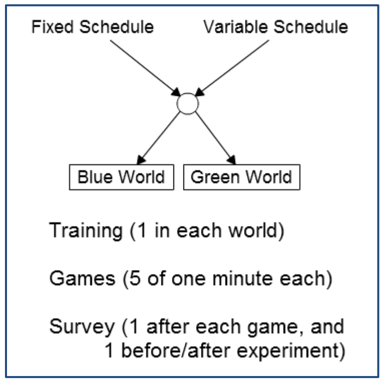

Figure 6: Experiment flow

After training, participants played a series of 5 games (7 in later experiments). Before each game, participants were given the choice of which world to play in (as in Figure 5). Surveys after each game ask the participants to rate both their enjoyment of the last game played and whether they enjoyed the variable world more than the fixed world on a 5-point scale, while surveys before and after the experiment ask additional questions (see Appendix A).

Experiments were conducted using Amazon Mechanical Turk (MT), a crowdsourcing Internet marketplace. Requesters post HITs (Human Intelligence Tasks), such as tagging a series of images. Workers browse for tasks and complete them for a monetary payment set by the Requester. Requesters can set conditions before a Worker is allowed to start a task, such as minimum approval rate. The only condition we set for our experiments was a minimum approval rate of 95% (the ratio of number of approved tasks to the total number of tasks completed). This was used as a basic quality control filter.

Our Mechanical Turk HIT contained a link to our game, which was hosted on Princeton servers and created using Adobe Flash. Game data was automatically sent via a PHP script to a MySQL database once the participant completed the Flash portion of the experiment (this includes the in-game and pre-experiment survey questions). Participants then completed the rest of the Mechanical Turk HIT (post-experiment survey) to finish the experiment. Data was matched between the Mechanical Turk HITs and the Princeton MySQL database through a common subject ID generated in the Mechanical Turk HIT (which the user entered into the Flash game).

Our central hypothesis was two-fold:

1) Reward schedules influence game preference.

2) Variable schedules are preferred over fixed schedules.

The former follows from our belief that reward schedules are a driving force behind why we play games. The latter follows from an intuition that variability is exciting.

To test our hypothesis, we first analyzed game choices. Specifically, we compared how many times participants picked one reward schedule over the other. Next, we analyzed the post-game surveys in which we had asked participants how much they enjoyed each game (Figure 7). Finally, we analyzed responses to pre-experiment and post-experiment questionnaires (see Appendix A). In one experiment, we asked participants what their two favorite video games were. We then categorized each game into a genre by using the first category listed on the game's Wikipedia page. Games that couldn't be found online were discarded.

Figure 7: Post-game Likert scale survey

To test for the significance of participant game choices, we ran a one-sample t-test with the null hypothesis of equal numbers of variable and fixed games played. In each experiment, we calculated the Pearson linear correlation coefficient between the number of variables games played by participants and their ratings for the variable world. We verified that this correlation was statistically significant in every experiment as a sanity check. Finally, we ran a one-sample t-test with the null hypothesis that participant ratings for the variable world and fixed world were identical. If a participant played only one world throughout the experiment, that participant had no ratings for the other world and was ignored for this test. We used an alpha level of .05 for all statistical tests.

In each experiment we contrasted a fixed reward schedule with a variable reward schedule. The reward consisted of points received from killing a monster. The fixed schedule remained the same in all experiments: all monsters gave a fixed 10 points. We compared this fixed reward schedule with a total of five variable reward schedules:

1) 5-15 Uniform Distribution (N = 173)
2) 0-20 Uniform Distribution (N = 186)
3) 5, 15 Binary 50/50 Distribution (N = 707)
4) 0, 20 Binary 50/50 Distribution (N = 196)

5) 0, 50 Binary 80/20 Distribution (N = 197)

Notice that in each distribution, the mean is the same as the fixed reward schedule, 10 points. During gameplay, we counterbalanced for the mean in the variable reward schedule by ensuring 10 points on average per monster after multiples of ten monsters were killed. That is, the last five monsters in each set of 10 balanced the first five (so if the first monster gave 6 points, the sixth monster would give 14 points). There exists a trade-off between having the variable reward schedule feel random yet keeping the range of possible total points bounded, and we felt this choice to be reasonable. This balancing for the mean occurred across multiple games up until the end of the experiment.

The Binary 5, 15 experiment (3 above) has almost four times as many participants as the other experiments (N = 707). This was because we ran four separate experiments, the first of which used a five game experiment, the last of three which used a seven game experiment. In one of the seven game experiments, we asked participants what their two favorite video games were. In another, we lowered the game timer from 60 seconds to 40 seconds. The switch to seven games and lowering of the game timer was to test the effect of these factors on game choices, which we found to be negligible. We will refer to Binary 5, 15 as the aggregate statistics of the three seven game experiments (we believe their similarity warrants this) except in two cases: the first is when we look at game genre and reward schedule preference (this refers only to the experiment in which we ask for participants' favorite video games) and the second is when we analyze participant explanations for preference of a reward schedule (this refers to all four experiments).

RESULTS

Variation is Preferred

Contrasting the 5-15 Uniform distribution with the fixed schedule of 10 points showed that participants chose to play the variable world more often ($t_{158} = 3.46$, $p < 0.001$). On average, a participant chose the variable world 2.87 times per experiment and the fixed world only 2.13 times. Contrasting the 5, 15 Binary distribution with the fixed schedule showed, again, that participants preferred the variable schedule ($t_{459} = 2.01$, $p < 0.05$).

These results support our hypothesis that reward schedule influences game preference, and that variable reward schedules are preferred. As we'll see in the next section however, the range of the variable schedule is a critical component in its preference or lack thereof.

Aversion to Smaller Rewards

Interestingly, increasing the range of the variable rewards from 5-15 to a 0-20 Uniform distribution caused the preference for the variable world to disappear ($t_{184} = 0.09$, $p = 0.93$). Similarly, changing the Binary distribution from 5, 15 to 0, 20 caused preferences to shift to the fixed schedule. However, this result was not manifested in the choice proportions per se, but rather is based on survey answers (participants rated the fixed schedule more enjoyable, $t_{160} = -3.40$, $p < 0.001$). Note that in all experiments, the correlation between survey answers and number of variable games played was significant ([5-15 Uniform, correlation = 0.55, p < 0.001], [0-20 Uniform, correlation = 0.54, p < 0.001], [5, 15 Binary, correlation = 0.43, p < 0.001], [0, 20 Binary, correlation = 0.70, p < 0.001], [0, 50 Binary, correlation = 0.46, p < 0.001]).

We carried out a final experiment using a 0, 50 Binary distribution with the participant scoring 50 points roughly 20% of the time. In this experiment, participants again preferred the fixed schedule, both in terms of game choices ($t_{173} = -2.72$, $p < 0.01$) and survey answers ($t_{162} = -3.61$, $p < 0.001$).

Individual differences in reward preferences

As part of the 5, 15 Binary experiment (in which participants preferred the variable schedule), we also asked participants what their two favorite video games were in a pre-experiment survey. In analyzing these data, we considered only those participants who strongly favored either the variable or fixed worlds. We defined strong preference as choosing to play one world at least two times more than the other. For these participants we found the following distribution of preferences across game genres:

Game Genre	Strongly Prefer Variable	Strongly Prefer Fixed
adventure	0.00%	1.19%
sports	10.08%	11.90%
run and gun	2.52%	2.38%
fighting	2.52%	1.19%
action roleplay	8.40%	2.38%
shooter	19.33%	14.29%
roleplaying	10.92%	5.95%
platformer	11.76%	23.81%
racing	12.61%	11.90%
puzzle	3.36%	8.33%
action-adventure	10.08%	8.33%
strategy	5.88%	7.14%
party games	2.52%	0.00%
poker	0.00%	1.19%
total	**119**	**84**

Table 1: Game genre and reward schedule preference. Percentages in each column add up to 100% (i.e., these are percentages of the total participants who preferred one type of reward schedule). The bottom row totals the number of participants in each group.

Overall, more participants strongly preferred the Binary 5, 15 variable reward schedule (N=119, compared to N=84 who preferred the fixed reward schedule). The percentages are percentages of the total in that column. There are apparent discrepancies in preference for some game categories. People who preferred the variable schedule were almost four times more likely to list an action roleplaying game than people who preferred the fixed reward schedule. This is reasonable to intuit. Action roleplaying games such as Diablo 2 are littered with variable reward schedules, for example, the items dropped by monsters are randomly generated from a probability table. Alternatively, people who preferred fixed reward schedules were twice as likely to list a platformer game as people who preferred variable reward schedules. This is also intuitive; a platformer game such as Mario Brothers offers little in the way of randomness as the rewards in each level are fixed at preset locations.

<u>Benefit of the Variation</u>

Analyzing 1459 Mechanical Turk post-experiment surveys, we looked in particular at what participants said in response to "If you chose to play one world more often than the other, why do you think that was?" We found that there were four categories of responses to this question:

1. *Participant prefers one world for <u>more points</u> – "In green world there is a factor of luck that might give you more points…"*
2. *Participants prefer one world for either <u>variety or steadiness</u> – "Blue because it allows for constant predictable progression…"*

3. *Participants prefer one world for purely <u>aesthetic</u> reasons – "Green is more pleasing to the eye…"*
4. *Participants prefer one world due to <u>aversion</u> towards the other – "In green world, fights against monsters that result in no experience feel pointless"*

Our findings are summarized in the table below. Symbol "v" is shorthand for variable world; "f" is fixed world, "g" for green, "b" for blue. Note that in the case of aversion (4 above) preference was always for the fixed reward schedule (the variable reward schedule was aversive).

	More Points	Variety/Steady	Aesthetic	Aversion	Participant Pool
Uniform 5-15	17v,7f	20v,5f	14g,2b	None	173
Uniform 0-20	5v,5f	18v,14f	9g,1b	5	186
Bernoulli 5,15	21v,8f	67v,34f	40g,6b	10	707
Bernoulli 0,20	14v,8f	13v,24f	15g,1b	10	196
Bernoulli 0,50	8v,6f	14v,29f	9g,1b	8	197

Table 2: Participant explanations for preference of reward schedule

The first "More Points" column shows that in all schedules more participants thought the variable reward games were more rewarding than the fixed reward games (except in Uniform 0-20, where an equal number of participants believed that each reward schedule is more rewarding). The average points de facto per monster in the variable world for each experiment were 9.91, 9.84, 9.85, 9.93, 9.88 for the Uniform 5-15, Uniform 0-20, Bernoulli 5, 15, Bernoulli 0, 20, Bernoulli 0, 50 schedules respectively. That is, in all cases the fixed schedule (that rewarded with 10 points per monster) was, on average, the more highly rewarding schedule.

Checking the specific point averages scored by those participants citing more points for the variable world in the Bernoulli 5,15 case reveals that roughly half of the participants scored more than 10 points, and the other half scored less than 10 points on average. Therefore, the belief that in one world one could score more points than in another was not necessarily grounded in reality. However, mostly all of the participants who thought they had a better chance of scoring more points in the fixed world actually scored more points in the fixed world. These are participants that had a bad initial experience in the variable world and switched over to the fixed world.

DISCUSSION

We analyzed 1459 participant choices and survey answers in a custom-built video game to determine reward schedule preferences. We discovered that participants had a preference for variable reward schedules (5-15 Uniform and 5, 15 Binary schedules were both preferred over a fixed 10-point schedule), except when the variable schedule's range was extended (5-15 Uniform was extended to 0-20 Uniform and preference disappeared; the fixed 10-point schedule was preferred over the 0, 20 Binary and 0, 50 Binary schedules). Furthermore, our results suggest that reward schedule preference can be an indicator for game genre preference.

Initially it had seemed that variable reward schedules are preferred in general (5-15 Uniform and 5, 15 Binary), however, extending of the variable schedule caused participants to shift towards the fixed reward schedule. It is not clear from these experiments alone whether the possibility of zero and near-zero rewards contribute to this shift, or whether it is wholly attributed to the increased range in the schedules. If we shifted the mean of all of our schedules upwards by 10 points, would we still see a preference for a fixed reward schedule of 20 points over a 10, 30 Binary distribution (originally a 0, 20 Binary)? We speculate that zero or near-zero rewards are intrinsically dissatisfying, but additional study is required to verify this. The shift towards the fixed reward schedule with increased range in the variable reward schedule suggests the existence of a "sweet spot" in variability.

There exists an apparent discrepancy in reward schedule preferences among different game genres, which can be exploited to recommend games with reward schedules similar to those of previously played games. While we have used game genres here, it is sometimes the case that games in the same genre have different reward schedules. Characterizing each specific game's reward schedules would give us enormous power in predicting new games that a player would enjoy. Moreover, designers of a new game would do well to be aware of reward schedules expected by players of the genre, trading the risk of potentially alienating their target audience with attracting players from other genres. It may also be of consequence to analyze how reward schedule preferences vary across the different player archetypes described in Yee (2006) [9].

These results can be seen through the lens of other literature, such as prospect theory, where people evaluate gains and losses from some status quo reference point [11]. According to prospect theory, the value function for losses is steeper than that for gains, so losses "loom larger" than gains (i.e. losses hurt more than gains gratify). This research, in conjunction with animal conditioning theories, provides a backdrop for understanding our results. Prospect theory serves as a possible explanation for why extending the range of the variable schedule had a negative effect on preference for that schedule.

The empirical use of these variable reward schedules is readily apparent in slot machines. Based on the results we obtained (that suggest a possible aversion to smaller rewards), we propose a reward schedule that caters to more people. Instead of giving zero reward when one does not win at a slot machine, the machine might reward with a variable number of points per play (redeemable for drinks or food) while only slightly lowering the jackpot values to compensate. In this way, players will feel that they are always earning something for playing.

Understanding the reward preferences of players can also impact the personalization of game-play. The idea of a personalized game experience has been around for quite some time, and it has often manifested itself in varied difficulty levels. Togelius (2007) created a racing game which learned the player’s racing style, then evolved completely new racing tracks optimized to the player’s skill level [12].

Analogous to dynamically adjusting difficulty, one can dynamically adjust reward schedules as more information is learned about the player. Just as over time people require more and more difficult racing tracks (paced differently for each player), people may require increasingly varied and novel reward schedules (again, paced differently for each player).

In the same vein, we can form better player models using reward schedule preferences. The ubiquity of virtual worlds heralds an age in which vast amounts of data are collected about large numbers of people. Kaminsky (2008) trained machine learning models on players doing mouse exercises, playing StarCraft, and playing Solitaire [13]. From mouse data features such as click speed and click length, the exact participant playing (from a pool of 15 people) could be determined using new StarCraft test data, with 80% accuracy.

Similarly, we collected data on players’ actions within games, such as which reward schedules they preferred. With better player characterization, we can potentially build better game recommendation engines, as well as detect unauthorized users in an increasingly virtualized world.

CONCLUSION

The goal of this work was to test whether reward schedules affect a player's preference of game, and whether variable reward schedules are preferred over fixed reward schedules (both of which are supported by our results). To achieve this, we created a custom-built video game in Adobe Flash and contrasted five different variable reward schedules with a single fixed schedule. We then analyzed both the quantitative and attitudinal data of 1459 Mechanical Turk participants.

In doing so, we found that variable reward schedules were preferred (5-15 Uniform and 5, 15 Binary schedules were both preferred over a fixed 10-point schedule), except when the schedule's range was extended (the fixed 10-point schedule was preferred over the 0, 20 Binary and 0, 50 Binary schedules, and there was no preference for either schedule when using a variable reward schedule of 0-20 Uniform). Furthermore, our results suggest that reward schedule preference can be an indicator for game genre preference.

Armed with this knowledge, we can begin to optimize reward schedules by using a variable schedule and varying its range to find the ideal variability. We should avoid small or non-existent (zero) rewards. Moreover, reward schedule preferences can be used to recommend new games, even across different game genres, with similar reward schedules. While reward schedules tell only part of the story of why we play one game over another, it is a critical piece of the puzzle that will become increasingly so as game-like mechanics all around us compete for our attention.

The abuse of video games has caused countless miseries such as addiction and neglect. But they've also brought people together to enjoy, compete, and celebrate. We often ask how we can make something better. Perhaps the deeper, more important question here, isn't how can we make better games; but how can we find parallels in the real world, so that we can apply what we've learned from video games to make the real world a more inspiring, and motivating place, to live.

REFERENCES

1. Ahn, L.v. and L. Dabbish, *Designing Games With a Purpose.* Communications of the ACM, 2008. 51(8): p. 58-67.

2. Popplestone, J.A. and M.W. McPherson, *An illustrated history of American psychology*. 2nd ed. 1999, Akron, Ohio: University of Akron Press. xiv, 222 p.

3. Herrnstein, R.J., *Aperiodicity as a factor of choice.* Journal of the Experimental Analysis of Behavior, 1964. 7(2).

4. Catania, A.C. and G.S. Reynolds, *A quantitative analysis of the responding maintained by interval schedules of reinforcement.* J Exp Anal Behav., 1968. 11(3 Pt 2 Suppl): p. 327-383.

5. Zeiler, M.D., *Fixed and variable schedules of response-independent reinforcement.* J Exp Anal Behav., 1968. 11(4): p. 405-414.

6. Mazur, J., *Fixed and variable ratios and delays: further tests of an equivalence rule.* J Exp Psychol Anim Behav Process., 1986. 12(2): p. 116-24.

7. Bateson, M. and A. Kacelnik, *Preferences for fixed and variable food sources: variability in amount and delay.* J Exp Anal Behav., 1995. 63(3): p. 313-329.

8. Bartle, R. *Hearts, Clubs, Diamonds, Spades: Players Who suit MUDs*. 1996.

9. Yee, N., *Motivations of Play in Online Games.* CyberPsychology and Behavior, 2006. 9: p. 772-775.

10. Romero, R., *Successful Instrumentation (Tracking Attitudes and Behaviors to Improve Games).* Game Developer's Conference, 2008.

11. Kahneman, D. and A. Tversky, *Prospect Theory: An Analysis of Decision under Risk.* Econometrica, 1979. XLVII: p. 263-291.

12. Togelius, J., R.D. Nardi, and S.M. Lucas, *Towards automatic personalised content creation for racing games.* Proceedings of IEEE Symposium on Computational Intelligence and Games (CIG), 2007: p. 252-259.

13. Kaminsky, R., M. Enev, and E. Andersen *Identifying Game Players with Mouse Biometrics*. 2008.

APPENDIX A

Mechanical Turk Survey:

What did you think the difference between the Blue/Green worlds was?

If you chose to play one world more often than the other, why do you think that was?

If you were to download this game to play on your personal computer as much as you liked, but could only play in one of the two worlds (i.e. once you choose to play in the Blue world, you will not ever have the chance to play the Green world and vice versa), which world would you choose?

Did you run into any problems while playing the flash game?

Pre-experiment Survey:

Subject ID: _______
Favorite Video Game: _________
Second Favorite Video Game: _________
Age: _______
Gender: _______
Annual Income (US$): _____________
Race: _____________
Education Level: _____________

Male/Female

White / African American / Asian / American Indian and Alaska Native / Native Hawaiian and Pacific Islander / Other

Middle School / High School / College / Bachelor / Masters / PhD

In-game Survey:

I enjoyed playing the last game (Green World):

Strongly Disagree	Disagree	Neither agree nor disagree	Agree	Strongly Agree
○	○	○	○	○

I enjoy playing the Blue World more than the Green World:

Strongly Disagree	Disagree	Neither agree nor disagree	Agree	Strongly Agree
○	○	○	○	○

Note that in the above, the first question is asking about the last world played, and the second question always asks whether the participant prefers the variable world more than the fixed world.