Gaming Under the Influence: The Effects of Social Comparison on Task Performance

Jeremy Nguyen
Dept. of Electrical Engineering and Computer Science
York University
Toronto, Ontario, Canada M3J 1P3
nguye688@my.yorku.ca

ABSTRACT
Video game analytics has become a tool in which gamers can use to improve their performance. In this study, we compare the effects of social comparison on task performance using a leaderboard availability from the game Lunar Lander. Results indicate that participants with leaderboard accessibility had a slightly higher score than those without. The mean game score for participants with leaderboards was 227.8 and 206.1 without. Feedback after the experiment suggested that the game was difficulty to play and a learning curve was required.

Keywords
Analytics, influence, performance, social comparison, queue dodging, self-evaluation

INTRODUCTION

With the rise in the popularity of competitive gaming, resources for video game analytics has greatly increased. From popular games such as *League of Legends* and *Overwatch*, players are constantly seeking a competitive advantage, and access to this resource has become a way in which players can improve their play.

Social comparison, a theory proposed by social psychologist Leon Festinger [4], is the idea that there is a drive within individuals to gain accurate self-assessment. People are always looking for an accurate self-assessment and a comparable provides an assessment that can be referenced. It is this evaluation that lets one understand the possibilities of their potential; instigating the need for self-improvement [3]. From domains such as sports, comparing one's self to other athletes, emerges an important exercise in the context of acquiring contracts. Athletes are always looking to get paid, and a comparable oftentimes provide a reference point – via statistics – that players can use to argue for money.

And in the context of video games, video game analytics has provided gamers a new way to discover comparable players such that a better self-evaluation can be made. Games such as *League of Legends* now have dedicated websites revealing empirical data on fellow peers, giving players a resource in which they can evaluate win conditions. The term *queue dodging* has become a popular trend in gaming culture, defining the action of leaving a game before its start; due to its unfavourable win conditions analytically. Players dodge a queue because a queue dodge results in less points being lost in the ranking system than a loss. OP.GG (www.op.gg), a website that collects *League of Legends* data, provides statistical analysis on player tendencies and background, is oftentimes used to determine whether one wants to dodge a queue. A queue dodge is often instigated when revelation of poor player tendencies on one's team or strong analytical data on the opposing team is revealed. People play games so they can win but seeing strong analytical stats of an opposing team or poor stats from one's own can be demoralizing. People make self-evaluations of themselves based on things that are attainable, and things that are not within limitation can be discouraging [4]. Players want to win and playing amongst people with their same skill level will help them improve. As such, given that all matchups are fair, the effects of one's performance when given analytical data is still largely unknown.

While this paper looks at video games in the context of social comparisons, it also looks at the intrinsic value gamers gain when looking up the analytics of their peers. Using a custom game application DemoLunarLanderPlus, we seek to understand the effects of in-game performance when shown statistical information of peers and how they differ when not.

Related Work

Colusso et al. [2] conducted a similar experiment to improve the efficacy of social comparison feedback. They conducted tests on two design strategies: (1) comparing users to a target described as a similarly experienced player and (2) adjusting the visual representation of performance so player scores appear closer to comparison target. The effects for social comparison on player performance were evaluated using the online game Flappy Bird, where the objective is to keep the bird from falling to the ground while avoiding obstacles (see Figure 1). In a controlled experiment, both feedback techniques improved game performance, but only coming from experienced players [2]. A possible explanation that these strategies mostly affected the experienced gamers is that experienced gamers cared more about the game. If someone cares about an activity, it is natural that they are more likely to be affected by it [2]. Inexperienced gamers meanwhile did not seem to be affected by the proposed design strategies. Inexperienced players were more interested in exploration and playing the game to discover things [2]. As such, our experiment takes into account that

our game is played by only inexperienced players and that the need for discovery is very minimal.

Zuckerman and Gal-Oz [5] performed a study evaluating the effectiveness of social comparison for promoting activity. Using a self-developed application called "StepByStep," they found that offering continuous measurement of walking time, a daily goal, and real-time feedback on progress toward a goal, greatly facilitated activity and increased walking time on average [5]. Gamified versions offering social comparisons revealed to be just as effective as the application [5]. However, due to contradicting findings from prior studies and the lack of research in this field, the assumption that gamification makes physical activity more enjoyable cannot be fully supported [5]. However, details of this study reveal analysis on how gamification can be used in our own experiment to increase motivational behaviour and improve overall performance.

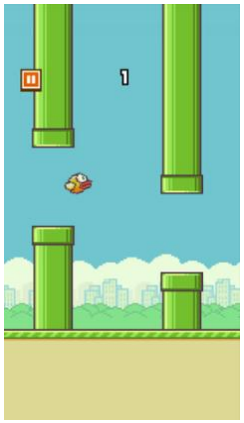

Figure 1. A screenshot of the game, Flappy Bird.

Blair [1] conducted a study that looked at the taxonomy of game design to improve performance, self-efficacy, and motivation in players. Looking at the influences of player behaviour and the guides of game design for achievements, it conducts three achievement tests: (1) expected, (2) unexpected, and (3) incremental where notifications before and after achievements are earned. His research showed that individual game mechanisms had little effect on players whereas multiple game mechanisms showed significant improvements in areas such as completing easy tasks, player attitudes, intrinsic motivation, and knowledge organization. The experiment correctly predicted multiple outcomes, however, mean scores from the knowledge performance trials were too high and unusable.

System Description

Participants with prior experience were rejected to ensure that skilled players did not set higher scores to discourage newer players.

A setup page was designed to gather participant data via the application (See Figure 2).

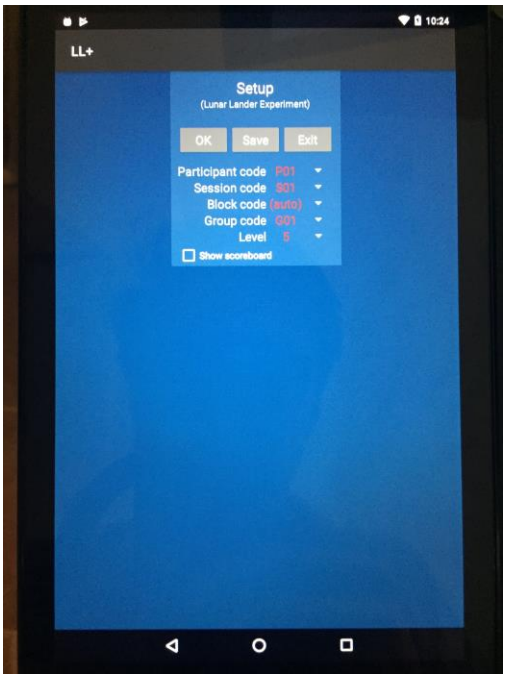

Figure 2. A view of the setup page.

Scores were recorded based on a multiplier on the amount of time and fuel used to land the spacecraft as follows:

- Landing score = 100
- Time score = 5 * (100 / completion time)
- Fuel score = 5 * (remaining fuel)

To develop a scoreboard, eight participants (four male, four female) played the game without use of a scoreboard. These participants had minimal experience and were used to set an initial scoreboard for subsequent participants.

The scoreboard is an option as part of the application's interface and was available to specific group types along with a setup page.

Game difficulty were increase as the levels progress:

- Less fuel available at higher level
- A faster landing is not allowed
- A greater lander angle is not allowed
- The landing pad is narrower

Once the game starts, the scoreboard is viewed by clicking on the settings button and the option for scoreboard is chosen (Figure 3). The scoreboard is sorted from highest score to lowest score (top to bottom) with only the top five scores displayed as shown in Figure 6. The scoreboard comprises of the eight initial scores recorded by the previous participants, where participants who score higher than one of the top five scores were replaced by a person on the scoreboard.

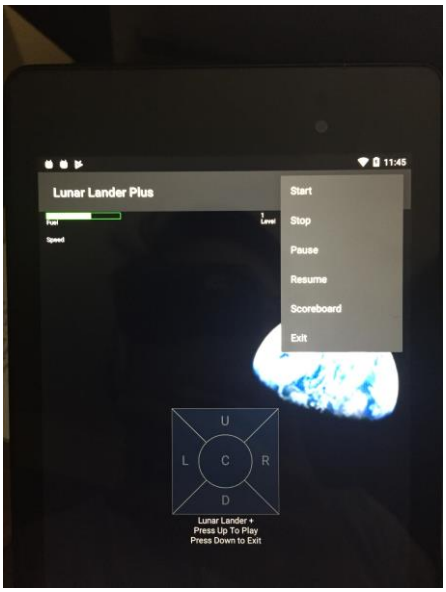

Figure 3. The scoreboard option as seen from participants with the scoreboard option via setup.

A post experiment interview was conducted to acquire feedback on the game overall.

METHOD

A user study was conducted to determine if participants who have a scoreboard will have a higher score than those without a scoreboard.

Participants

Twenty participants (10 male, 10 female) were recruited from York University. Ages ranged from 18 to 25. All participants had no previous experience with the game, Lunar Lander.

Apparatus

The experiment used Google's *Nexus 7* (2013) tablet device measuring 19.9 cm by 12.0 cm by 1.5 cm. The screen display is 1280 × 800 pixels and pixel density is 216 pixels/inch. The device runs on Android 5.0.1 Lollipop. See Figure 4.

The experiment ran a custom application of the video game Lunar Lander developed by Atari. Software developed in Java using Android Studio.

A quick interview will be conducted to be used to measure the level of the participants' familiarity or experience with the game.

Figure 4. Google's *Nexus 7 (2013)*

Procedure

All participants were asked to complete a quick interview on their experience with the game Lunar Lander. Participants with prior experience were rejected from continuing the experiment while participants with minimal experience will be accepted.

Participants were given a quick demonstration and allowed to play one level before starting.

Each participant was given five attempts (lives) to reach a high score, and an attempt is used up each time the player improperly lands the vehicle. See Figure 3.

Figure 5. A screenshot of when the player life is used up, as indicated by the red circle around the spaceship.

Participants were sorted into two groups: one without the option to view a scoreboard, and one with. Participants were given code names, and these code names represented their scores (Figure 6). They were split evenly as follows:

- Group 1 (no scoreboard): 3 males, females
- Group 2 (scoreboard): 3 males, 3 females

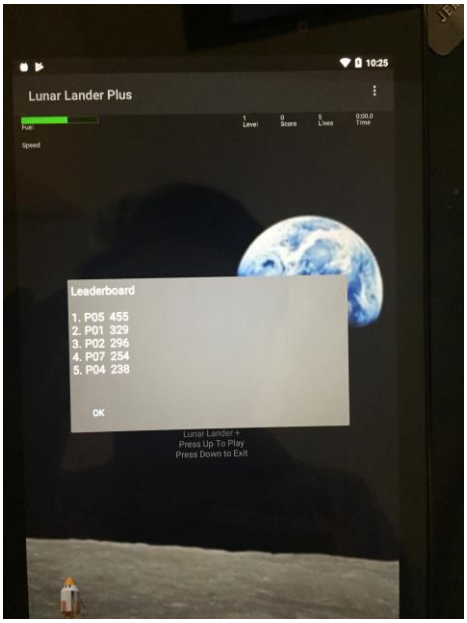

Figure 6. A view of the scoreboard.

Design

The experiment was a 2 × 5 mixed subject design where one group is given a scoreboard indicating the highest scores and another were not. There were two independent variables:

- Input method: participants with no scoreboard, participant with a scoreboard
- Life: 1, 2, 3, 4, 5

Participants used to create the initial scoreboard (eight) were excluded from the trials.

There were five attempts (trials), thereby offsetting any learning effects. Trials lasted as long as the participant did not use up all their attempts but went on no longer than 10 minutes; where participants were forced to end. Trials are based on the game's levels and got harder as the participant completes a level.

The dependent variables include the participant's completion time and in-game fuel. Completion time was measured in seconds up to two decimal place. In-game fuel and speed was be measured as imaginary floats determined by the game measured up to two decimal place.

The total numbers of trials was 12 participants × 2 input methods × 5 lives = 120 trials.

RESULTS AND DISCUSSION

Scores for participants with an active scoreboard and those without were similar, however participants with a scoreboard scored slightly higher. Participants with access to a scoreboard had score averages of 227.8 while participants without had score averages of 206.1 (See Figure 7). Based on the experiment conducted by Colusso et al. [2], the results of the participants with scoreboards were expected as game performance, in general, was improved. Participants who had access to a scoreboard had continuous measurement of their scores and were provided feedback on their progress towards completing the task; facilitating increased performance [6].

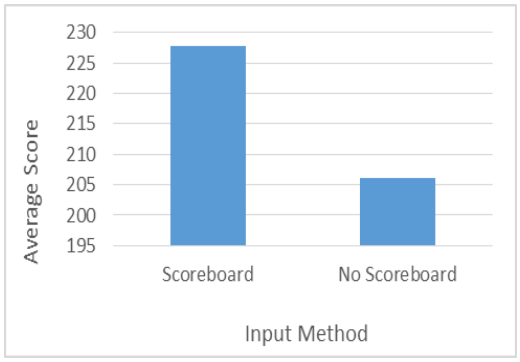

Figure 7. Average scores of each input method.

Completion times for both groups varied greatly overall. Means for level completions for participants without a scoreboard was 23.13 and 21.64 for participants with.

Average number of trials for participants with scoreboard was 7.2 trials per game and 7 trials per game without, suggested a difficulty in the game. Average levels for participants also confirm that players did not get far as results for average levels were 2.098 for players with a scoreboard and 1.374 for those without. As stated by the experiment by Colusso et al. [2], because of the inexperience, participants were not affected by the proposed design. Inexperienced participants were more interested in exploration and playing the game to discover things. Feedback from participants after game revealed they were more interested in learning the game and its mechanics. Some participants who held scoreboard indicated that they were much more focused on learning the game than beating someone else's high score.

CONCLUSION

In this user study, social comparison had improvement on scores and completion times for participants with a scoreboard. However, little improvement was observed from scoreboard revelation and without. Participants showed an interest in learning the game suggesting that the game was hard, or the participants gathered were not frequent gamers. The experiment did not take into account the overall gaming tendency but only experience of the game Lunar Lander.

Feedback from participants after the experiment revealed that the game was hard to play. Participants who claimed they were competitive via the pre-experiment interview had stated after that they spent most of the time learning to play the game than looking at the scoreboard to achieve a new high score.

Future Work

Future work includes modifying the pre-experiment interview to gather information on participants' gaming

history as frequent gamers show more competitiveness than inexperienced players [2].

Further work on Zuckerman and Gal-Oz's experiment can be done using a real-time scoreboard as opposed to a scoreboard that is available upon request.

REFERENCES

1. Blair, L. 2011. The use of video game achievements to enhance player performance, self-efficacy, and motivation. University of Central Florida. Orlando, Florida
2. Colusso, L., Hsieh, G., and Munson, S. A., 2016. Designing closeness to increase gamers' performance. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, 3020-3024. New York: ACM.
3. Dai, L., and Xiao, R. 2016. The influence of social comparison on job performance. *Open Journal of Social Sciences*. 4. 147-151. China.
4. Festinger, L. 1954. "A theory of social comparison processes". Human relations. 7 (2): 117-140.
5. Gino, F., and Staats, B.R. 2011. Driven by social comparisons: How feedback from coworkers' effort influences individual productivity. Harvard Business School Working Paper. 11. 78.
6. Zuckerman, O., and Gal-Oz, A., 2014. Deconstructing gamification: evaluating the effectiveness of continuous measurement, virtual rewards, and social comparison for promoting physical activity. *Personal and Ubiquitous Computing*. 18. 7. 1705-1719.