Mobile Gamification for Experiment Data Collection: Leveraging the Freemium Model

Kristen Dergousoff
Computer Science
University of Saskatchewan
Saskatoon, SK, S7N0Y7
kristin.dergousoff@usask.ca

Regan L. Mandryk
Computer Science
University of Saskatchewan
Saskatoon, SK, S7N0Y7
regan.mandryk@usask.ca

Abstract

Classic ways of gathering data on human behaviour are time-consuming, costly, and are subject to limited participant pools. Gamification provides a motivation to participate, but also requires the development of specialized, research-question specific games that can be costly to produce. Our solution leverages the popular Freemium model of play to motivate voluntary participation by rewarding players for participation in microexperiments with in-game powerups, using a robust framework to study multiple unrelated research questions within the same system. We deployed our game on the Android store and compared it to a gamified laboratory version and a non-gamified laboratory version, and found that players who used powerups were motivated to do the microexperiments.

Author Keywords

Gamification, freemium, psychophysics, crowdsourcing

Introduction

A fundamental, and limiting, step in Human Computer Interaction research is gathering data in order to understand human behaviour. Researchers often perform costly and time-consuming user studies in laboratory environments. In academic research this

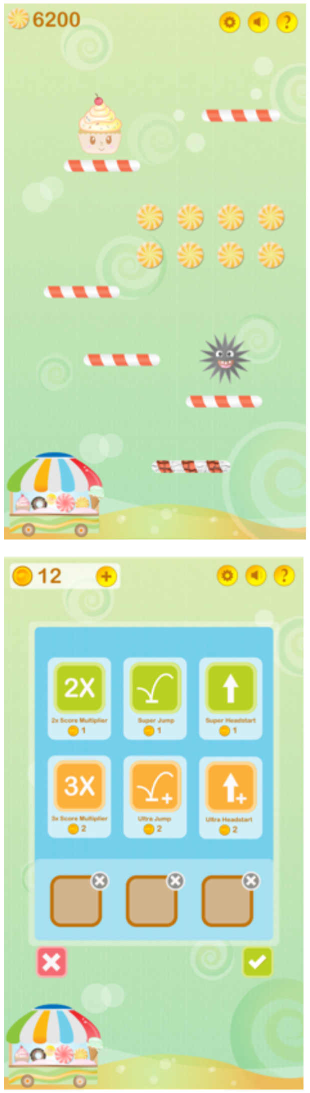

Figure 1. Top- Prototype of gameplay; Bottom- Prototype of powerup menu

usually consists of recruiting participants from a participant pool at the university, which confines participation in studies to the "boom-and-bust semester cycle" and limits the generalizability of the study to the demographic of young college students [9] [10].

One solution that has been proposed is to crowdsource experiments. Crowdsourcing platforms such as Amazon's Mechanical Turk (AMT) have been shown to successfully recreate experimental results [6], [7] and provide additional benefits such as constant any-time access to a large and diverse participant pool [9]; however, issues have been raised regarding data quality and the ethics of low pay [9].

The gamification [4] of experimental studies has been shown to motivate large numbers of voluntary participants [11] and to increase the enjoyability of tasks [5]. However, it is still unclear whether the results obtained through a gamified approach are as accurate as data gathered in traditional laboratory settings [1]. Also gamified experiments are usually highly customized to the specific research question, requiring costly development of new games for each research problem that is addressed.

Our solution is to combine the motivation of gamification with the broad participant base available through crowdsourcing by leveraging the benefits of the "Freemium" model, in which players can play a game for free, but are given access to special content, features, or advertisement-free play for a fee. We created a framework in which players of a game gain in-game advantages (i.e., powerups) for completing experimental tasks (i.e., microexperiments) [3]. We target the mobile game market to target players looking to kill time with short gameplay sessions. The main advantage of our framework is that it separates the system for completing microexperiments from the game so that multiple experimental tasks can be deployed in a single game, or multiple games can be deployed to increase the appeal for players. To evaluate the efficacy of our framework, we developed one game and two experimental tasks.

System

Our game, Sugar Rush, is an action-based vertical platformer that utilizes the accelerometer on an Android phone. In the game, players are in control of a continuously bouncing cupcake, which they guide through platforms and enemies to collect candy and coins and obtain a high score (see **Figure 1**). Players are provided with the option to purchase useful powerups with in-game currency earned through the completion of quick microexperiment tasks.

Powerup items can be accessed from the main menu (see **Figure 1**) and "purchased" by players using in-game credits. These powerups provide players with a benefit to gameplay, such as increasing the value of each candy collected, or blasting the player through the first sections of the game, collecting a large amount of coins along the way. Once powerups have been purchased, they are automatically applied to the next game session that the player starts.

Players can earn in-game credits to purchase powerups by completing the quick experimental tasks that are included in the task manager. When the player chooses to participate in a microexperiment, the system randomly chooses one of the available experimental tasks. Upon the completion of the task, the player is

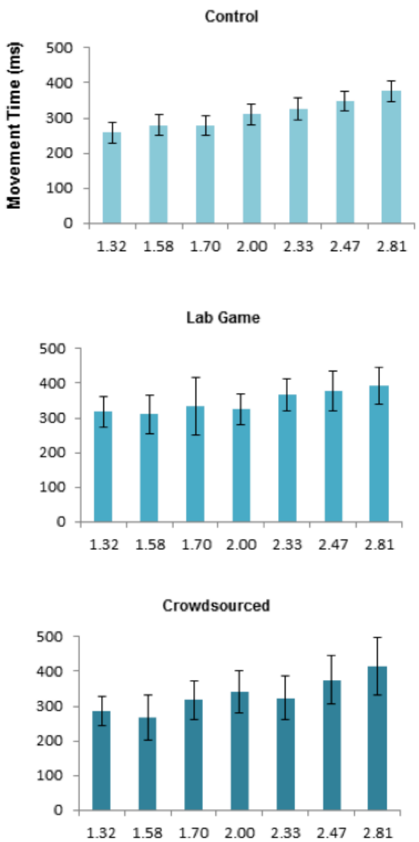

Figure 2. Mean movement time for the motor (Fitts) task for each index of difficulty across all three conditions.

awarded a single credit. The number of credits required to purchase a powerup varies between one and two credits and a player can purchase a total of three powerups at a time, providing incentives for the player to participate in multiple experimental tasks.

Evaluation

We evaluated our framework under three conditions to isolate the factors of gamification and crowdsourcing and examine their effects separately on the quality of data gathered (Crowdsourced Game, Laboratory Game, and Control, i.e., laboratory no game). We also analyzed usage data received from the Android market to inform the degree to which our prototype game and chosen Freemium model motivated voluntary participation in the tasks. Finally, we deployed a survey with the laboratory participants who played the game to gather their opinions on the topic.

We were interested in answering two main questions. First, whether the quality of the work that they are doing meets the standards set by similar experiments conducted using more traditional approaches, such as in the lab. Second, whether participants are motivated to play our game and whether they find enough value in the powerups to "work" for them by doing tasks.

Results and Discussion

Data Quality

Results of the evaluation of data quality indicated that there was no difference in performance between the three conditions for the performance-based motor task (i.e., a Fitts reciprocal tapping task [8]) – see **Figure 2**; however, the crowdsourced game resulted in worse performance on the attention-based cognitive task (i.e., a Cleveland and McGill information visualization judgment task [2]) – see **Figure 3**.

One possible explanation for this difference could be that participants are concerned with completing the task quickly in order to earn their credit. Because this concern of earning the credit quickly is in line with the performance-based nature of the motor task, it does not affect the quality of the results. However, because the cognitive task requires that participants take the time to make a judgment regarding the values they input, this goal of finishing as quickly as possible is detrimental to quality.

One way to motivate greater effort and accuracy by the participants in their responses is to make the rewards contingent on the apparent effort put into the task. If the data that is entered is easily identifiable as being false data, the credit reward can be withheld. Although this would only deter players from providing obviously false data, our results showed that when participants took time to enter a value, the value entered was not unreasonable. In addition, the value of the reward could be tied to the accuracy of the data, motivating participants to input quality data. Withholding rewards for obvious lack of effort or tying the value of the reward to the quality of the answer would encourage participants to spend more time in giving a response and likely result in better quality data.

Motivation to Participate

The survey results of player motivation suggested that players were willing to do the tasks in return for earned in-game bonuses, would prefer the tasks over in-game advertising, and that the tasks did not detract from the play experience. The usage data shows that participants who were exposed to the powerups used

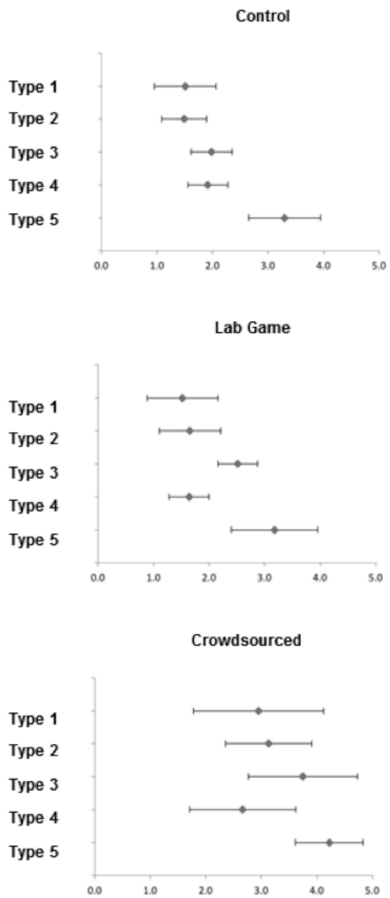

Figure 3. Mean log absolute error for the cognitive (McGill) task for each of five types of charts across all three conditions

them in about 21% of games, but that the majority of participants did not actually try the powerups.

We feel that it is possible to increase the motivation of players to complete experimental tasks in a framework such as ours through game design. Specifically, exposing the players to powerups in a limited manner would help players learn to appreciate the value of powerups (e.g., as is done in Candy Crush or Frozen's Free Fall) and likely increase the appeal. Hindering progress by unlocking content through microexperiment completion could also motivate increased participation.

Conclusions

Gathering data on human behaviour is a limiting factor in behavioural research. Crowdsourcing and gamification provide two possible solutions to this data collection problem but they each pose issues of their own. In this paper we presented our solution, which combines both gamification and crowdsourcing techniques into a smartphone-based platform to motivate voluntary participation and provide researchers with a framework that can be used to investigate multiple research questions without the need to develop costly specialized games.

Results from our initial evaluation showed that the quality of the motor task data did not suffer; however, the data from the cognitive task was of lower quality. We feel that tying the reward to the quality of the data could improve data quality for attention-based tasks. Despite a low adoption of powerup usage, participants that were exposed to the experimental tasks were supportive of participating in return for in-game benefits.

References

[1] Cechanowicz, J., Gutwin, C., Brownell, B., and Goodfellow, L. Effects of Gamification on Participation and Data Quality in a Real-World Market Research Domain. *Gamification '13*, (2013) 58-65.

[2] Cleveland, W. S., and McGill, R. Graphical perception: Theory, experimentation, and application to the development of graphical methods. *J. Am. Statistical Association, 79*, (1984), 531-554.

[3] Dergousoff, K., and Mandryk, R. Mobile Gamification for Crowdsourcing Data Collection: Leveraging the Freemium Model. CHI 2015 (to appear).

[4] Deterding, S., Bjork, S., Nacke, L. E., Dixon, D., and Lawley, E. Designing Gamification: Creating Gameful and Playful Experiences. *CHI 2013*, ACM (2013), 3263-3266.

[5] Fatla, D. R., Gutwin, C., Nacke, L. E., Bateman, S., and Mandryk, R. L. Calibration Games: Making Calibration Tasks Enjoyable by Adding Motivating Game Elements. *UIST `11*, ACM (2011). 403-411.

[6] Heer, J., and Bostock, M. Crowdsourcing Graphical Perception: Using Mechanical Turk to Asses Visualization Design. *CHI 2010*, ACM (2010). 203-212.

[7] Komarov, S., Reinecke, K., and Gajos, K. Z. Crowdsourcing Performance Evaluations of User Interfaces. *CHI 2013,* ACM (2013).

[8] MacKenzie, S. I. Fitts' Law as a Research and Design Tool in Human-Computer Interaction. *Human-Computer Interaction, 7*, (1992). 91-139.

[9] Mason, W., and Suri, S. Conducting Behavioural Research on Amazon's Mechanical Turk. *Behavior Research Methods*, (2012). 1-23.

[10] Reinecke, K., and Gajos, K. LabintheWild: Conducting Large-Scale Online Experiments With Uncompensated Samples. *To appear in CSCW'15.*

[11] von Ahn, L., and Dabbish, L. Labeling Images with a Computer Game. *CHI '04*, ACM (2004). 319-326.