River Rescue: Identifying iNaturalist Animals via Gamification

Jaiden Lee
Millburn High School

December 22, 2022

ABSTRACT

There has been a recent trend in using *gamification* (that is using game design in non-game scenarios) to attract new citizen scientists and make citizen science tasks more enjoyable. Most of the current research in gamification makes use of points, badges, and leaderboards. However, not much research has been done on adding different forms of gameplay. Thus, this study investigates the possibility of using gameplay to encourage participation in citizen science. In this paper, I conducted an experiment to evaluate whether adding hyper casual gameplay to a citizen-science application can influence user engagement while classifying observations of animals from iNaturalist (an online, global biodiversity index). I developed 2 games under the name of *River Rescue*: one with a casual gameplay aspect and one without, with participants being randomly assigned to each group. Results indicate that, for this group of participants, adding a casual gameplay aspect did not significantly impact user engagement. However, based on participant suggestions and qualitative data collected, my findings suggest that future work might want to research fast-paced games with progression, making sure to avoid repetitive/tedious tasks and to minimize loading times.

1. INTRODUCTION

Recently, there have been many efforts at using “citizen science”—that is, tasks where non-scientist citizens contribute towards scientific research [1]. One such example of a citizen science task is iNaturalist. iNaturalist is a database of wildlife information in which the general public can upload images/sounds of flora and fauna (known as “observations”) and classify other individuals’ observations.

Despite the usefulness of citizen science, citizen science tasks face significant challenges, among them is maintaining user engagement. Citizen science tasks can often be tedious, repetitive, and offer seemingly little to no benefit to the participants themselves. In fact, the majority of participants make extremely small, sporadic contributions and oftentimes stop contributing soon after [7]. Thus, to address this issue, successful citizen science tasks should be enjoyable and interesting.

One way to create enjoyable citizen science tasks is through the usage of games. In this research, I am focused on the usage of game design elements in non-play scenarios, also known as *gamification* [6]. In terms of citizen science tasks, gamification refers to the combination of scientific tasks with game-elements, such as points and badges, to attract more users and engage participants.

The current research in gamification and games often focus on the techniques of adding points, badges, and leaderboards [3]. For example, researchers developed a mobile application that gathers plant data, was gamified through the addition of badges and a leaderboard system. Although the results showed that many individuals were motivated to use the gamified application, the results also pointed that most individuals would not go out of their way to use the gamified Biotracker. The majority of modern-day video games, however, do not rely solely on points, badges, and leaderboards—instead, commercial entertainment games heavily rely on story, gameplay, and interesting game mechanics. Recently, some people have begun to turn their attention towards these aspects of commercial entertainment games and away from the conventional approach of using points and badges. For example, some researchers have turned towards the usage of storylines in diegesis as an effective method of gamification. Beyond story-based games, however, many commercial entertainment games rely on their gameplay instead of their story, as seen by the popularity of obstacle-course based games like Fall Guys, hyper casual games like Helix Jump, and RPG games like Elden Ring.

Therefore, to address this gap in current research, I investigate the question: does adding a gameplay aspect to citizen science games impact user engagement when classifying images of animals?

2. BACKGROUND

2.1. Gamification

One of the most common definitions of gamification is “the use of game design elements in non-game contexts” [6]. Despite the usefulness of gamification, there is also a lot of controversy regarding the term gamification. Some individuals, such as Ian Bogost, have referred to gamification as “exploitationware” [3]. The reason for a lot of this criticism against gamification is because gamification is oftentimes reduced to solely the addition of points and badges, which is why some individuals, like Margaret Robertson, believe that gamification should be called “pointsification” [10]. The criticism goes by saying that points and badges aren’t actually key elements of “games”. Rather, they are a means of measuring progress and do not “produce an experience of interest, enlightenment, terror, fascination, hope, or any number of other sensations” [3]. Despite the criticisms, points-based gamification has been shown to be successful, such as Duolingo, Old Weather, and EyeWire [11]. However, this present research focuses on modifying the gameplay, rather than the usage of points, badges, and leaderboards.

2.2. Related Work

Gamification is not new. Many applications, such as Duolingo, have already been using gamification for years. Gamification applies beyond learning new skills, however—it can also be used for citizen science. A notable example of a game-based application to contribute towards citizen science is Foldit, a free-to-play protein folding computer game. Foldit treats these protein-folding tasks as “puzzles” and gives players a score based on their performance.

In another study, researchers developed a gamified application called Biotracker [4]. As mentioned above, Biotracker is a mobile citizen-science application, allowing individuals to collect information on plants. In the study, researchers added badges and leaderboards to test if they can motivate millennials to contribute towards citizen science or not. The results suggested that, although most individuals would probably not go out of their way to use a gamified (badges and leaderboards) citizen science app, some individuals would.

More similar to our research, also mentioned above, research has been done beyond points, badges, and leaderboards. One study compared the effectiveness of points-based gamification against story-based gamification [11]. The results showed that individuals preferred using the story-based gamified application, Forgotten Island, over the points-based gamified application, Happy Match. Many individuals cited that the story-based game led to greater motivation and incentive as compared to solely points.

2.3. iNaturalist

iNaturalist is a community of citizen scientists, biologists, and naturalists. iNaturalist allows for individuals to upload observations of wildlife (including images and sounds) to a database. In addition to images and sounds, observations may also contain other data, such as location, sex, and whether the specimen is alive or dead. Once an observation is uploaded to iNaturalist, anyone can help to identify the observation. Individuals submit votes on what they think an observation is. Once two-thirds of voters agree on the taxon (i.e. species, class, family, genus) is, the observation is then considered to be “research-grade”.

3. METHODS

3.1. Participants

I recruited 8 participants from the United States. To find these participants, I used a combination of convenience sampling and snowball sampling: I reached out participants known to the researcher and asked them if they had any recommendations for other people to contact. Participants ranged in age from 12 to 17 years old (M = 16.0, SD = 18.7; 7 male, 1 female). No additional demographic data were collected.

3.2. Protocol

Before the study, participants were prompted with an informed consent form to be signed, which provided basic information about the study as well as the risks involved in the study.

To collect the data, participants were randomly assigned to either the gameplay group or non-gameplay group. 5 participants were randomly assigned to the gameplay group while the other 3 were assigned to the non-gameplay group. Then, each participant was sent a download link for their respected version of the game as well as a survey to fill out after they had finished playing the game for a minimum of 10 minutes. The survey was designed to gather both qualitative and quantitative data about opinions in regards to either version of the game.

Questions asked included level of enjoyment, interest in nature and video games, and qualitative opinion based questions, such as favorite part of the app.

In addition to the survey, data was collected during gameplay, which reported information on how long users actually played for and accuracy. Accuracy was only calculated for identifications that were already deemed research-grade (at least two-thirds of identifiers agreed on the taxon) and at least 3 separate users made identifications. This was to ensure that the taxon agreed upon already had consensus (fewer than three votes is not enough to come to a consensus).

3.1. River Rescue

I developed 2 games using the Unity game engine, both under the name of *River Rescue*. One version of the game contained a hypercasual gameplay aspect (G) and the other version did not (NG), to which participants were randomly assigned to play. The task at hand in each game was to classify images of animals. In the game, users would be prompted with a screen that displayed an original image of an animal with 2-4 images of similar looking species (Fig. 1.).

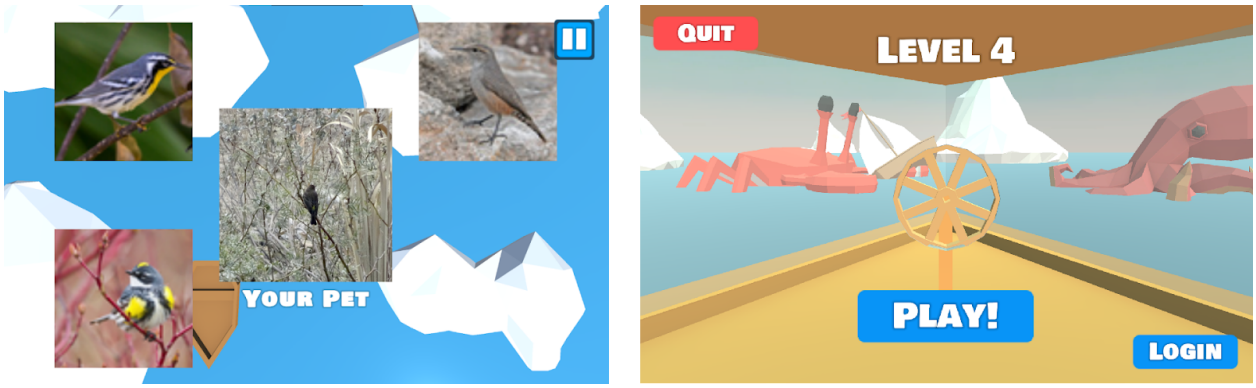

Fig. 1. *Screenshots of River Rescue*. Left: the menu to classify images of animals. The central image is the observation, and the surrounding images are image buttons that contain similar looking species of animals. Right: the main menu for *River Rescue*.

This task was framed around saving your kidnapped pet. If the user was logged in, choosing any one of these animal images would send an identification submission to iNaturalist's database. This is where the version without a gameplay aspect ended. For the NG version, the game would return to the main menu, and the user would be able to classify more animals again (See Fig. 3.).

The version with the gameplay aspect was modeled based on the hyper casual game genre. The player would control a ship that constantly moves side to side, and when the player clicks

the screen, the ship would move forward. The goal was to dodge all the obstacles in the way to finally “save” your kidnapped pet. While moving forwards, the user's powerup bar would charge up, which would allow the user to destroy all obstacles in their path for its duration (Fig. 2). After the level has been completed, the game would repeat (it would return to the main screen and the user would be able to play again) (See Fig. 3.).

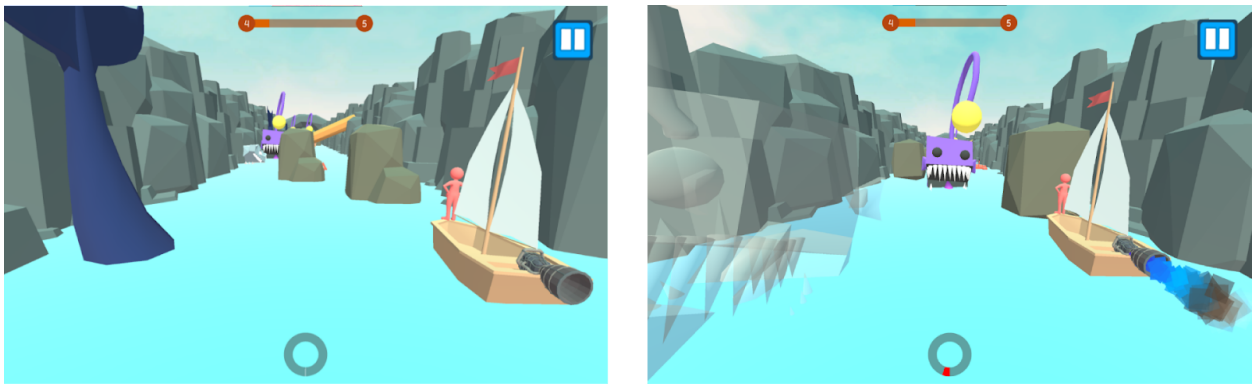

Fig. 2. *Screenshots of River Rescue Gameful Version*. Left: gameplay of the ship moving normally. If the ship crashes into any obstacles, the game will end. Right: powerup is activated (denoted by the blue fire and the red bar on the bottom); in this state, users can sail through obstacles and destroy them.

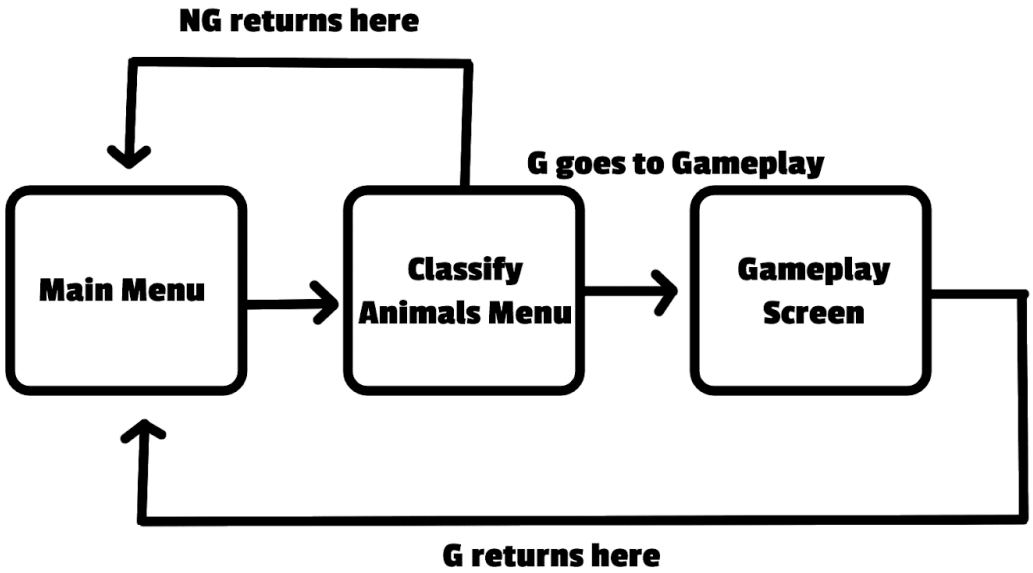

Fig. 2. *Gameplay Flow Diagram*. Players load in on the main menu. When they press play, the menu switches to the classify animals menu. For the non-gameful group (NG), screen switches back to the main menu and the cycle repeats. For the gameful group (G), the game switches to the gameplay screen, which then returns to the main menu screen afterwards.

4. RESULTS

4.1. Gameplay Log Statistics

After conducting the study, I compared the number of identifications, accuracy, and time played between the 2 groups. For the number of identifications, my results indicated that significantly more identifications were made by users in the non-gameplay group. When comparing their number of identifications, NG participants (M=13.0, SD=2.00) compared to the G participants (M=7.40, SD=1.67) demonstrated significantly more identifications, $t(6)=4.29$, $p=.005$.

However, while comparing the accuracy and time played between the 2 treatment groups, results were not statistically significant. The difference in time played between the NG participants (M=7.45, SD=1.82) and the G participants (M=6.82, SD=3.48) was not statistically significant, $t(6)=.283$, $p=.787$. Similarly, the accuracy of the NG participants compared to the G participants was not statistically significant, $t(6)=.1892$, $p=.856$.

4.2. Questionnaire Statistics

After receiving the survey results, I compared participants' self enjoyment level. I originally asked participants to rate their experience on a scale from 1 to 7, with 1 being extremely boring and 7 being very fun. I ran a Wilcoxon Rank Sum test (because the difference between rankings is most likely different for each participant) on this rating. The enjoyment level rating between the G participants (M=3.60, SD=1.14) and the NG participants (M=4.00, G=1.73) was not statistically significant, with $p=.643$.

In addition, I also asked participants to rate their interest in nature from 1 to 7. For the gameplay group, there was a mean of 4.8 with a standard deviation of .837, while the non-gameplay group had a mean of 2.67 and a standard deviation of 1.53. I ran a Wilcoxon Rank Sum Test, and determined that the interest in nature while comparing the G group (M=4.8, SD=.837) and the NG group (M=2.67, SD=1.53) had $p=.0651$. Although this indicates that the results are not statistically significant, the p-value was very actually close to the .05 threshold, which does demonstrate that, on average, the people in the gameplay group tended to be more interested in nature.

I also asked users qualitative questions describing their experiences with both applications. For the gameplay group, 4 out of 5 participants stated that the game was repetitive, and based on how they elaborated, many felt bored and/or frustrated. In fact, one participant said that he

would have preferred there not to be a game as he preferred the task of classifying animals, and the game made the task frustrating. 3 out of 5 of the participants in the gameful version actually indicated that the main reason they felt frustrated and/or bored was due to the lack of a sense of progress. They indicated that they would have rather played an RPG, adventure game, or at least played a game with a set amount of levels with different environments and obstacles that get progressively more difficult (similar to platformer games). In addition, in both groups, 5 out of 8 participants indicated that they disliked the long loading time for the game. However, in both groups, 7 out of 8 participants indicated that they really liked the design of their respective applications (including music selection, game ideas, animations, and 3D models).

Furthermore, I asked participants about how often they play video games. 5 out of 8 participants said they played more than 10 hours of video games a week. 1 played 5-10 hours a week, another played 3-5 hours a week, and one played less than 1 hour a month. This indicates that the overwhelming majority of participants had a lot of “gaming” experience. In addition, most individuals played rather “sophisticated” games, including MOBA games, Deepwoken, adventure/story games, FPS games, and horror games. No individuals reported that they played casual/hyper casual games.

5. DISCUSSION

5.1. Interpretations

According to my results, the fun rating between the G and NG group had p=.643. This indicates there is no convincing evidence that adding a gameplay aspect significantly impacts user engagement when classifying animals, at least for the participants in this study. However, this does not mean that the usage of gameplay is always ineffective in every scenario. In fact, in a game called Forgotten Island, the usage of story-based gameplay actually proved to significantly increase user engagement [11]. In my study, it was clear that the gameplay version was not as well received, however, as demonstrated by the boredom/frustration many participants felt. It is possible that, with improvements to the game, the gameplay version would have been more well-received, as most individuals indicated they disliked the game due to its long loading times and repetitive nature. In addition, the majority of individuals indicated that they played a lot of video games and did not play casual games. Therefore, it is possible that the gameplay version would have been better received by a group of participants that played less video games and/or played casual games. Furthermore, results indicated that, in the gameplay group, many individuals were very interested in nature, suggesting that they would be keen to frequently contribute to citizen science (especially classifying animals). And according to prior research on citizen science, generally high contributors were more engaged through competitive social features [2]. Thus, it is also possible that adding a competitive points-based incentive could have increased user engagement (given the large interest in nature). Moreover, participants indicated that they would have preferred to use a game with a sense of progression because, without progression, games can feel repetitive and/or frustrating. This idea of using

progression to engage interest is similar to previous research, in which, according to a study, badges encouraged usage of a gamified service [8].

5.2. Limitations

This study had several limitations. First, the sample size was extremely low—I only had 8 participants. Originally, 13 people had signed up for the study, but only 8 actually followed through and carried out the study, which means that there could also have been a form of non-response bias. Moreover, the treatment groups were unevenly sized. The non-gameplay group had 3 participants, compared to the 5 participants for the gameplay group, which means that a participant in the non-gameplay group weighed more than a participant in the gameplay group. Moreover, a large part of my data was self-reported. Individual participants may each have had different definitions and/or ratings of fun, as this was an opinionated question.

5.3. Future Research

For future research, it would be best to use a much larger sample size, as there is a much higher likelihood of bias in small sample sizes. In addition, because there are so many various forms of gameplay (casual, adventure, story, FPS, obstacle-course), other forms of gameplay should be tested to gain a full understanding on whether adding gameplay actually has an effect on user engagement for citizen science tasks. Because my results indicated that participants would have preferred games with progression, future research should most likely specifically investigate the effectiveness of using games that provide users with a sense progression (such as RPG games, adventure games, and/or platformer games) so as to reduce the feeling of repetitiveness/boredom users experience. In addition, these games should be quick, meaning that nothing should take too long (whether that is tedious tasks/missions or loading times). Furthermore, different groups of participants should be tested. For example, participants in this study had a strong interest in RPG games, horror games, and FPS games. In the future, participants with more varied interests and/or demographics should be tested.

6. CONCLUSION

This paper investigated the impact of using casual gameplay to encourage participation in citizen-science. Two applications were developed under the name of *River Rescue*: one with a hyper casual gameplay aspect and one without. Results indicated that adding a casual gameplay aspect did not significantly impact user engagement. This does not mean that gameplay is ineffective, however. In fact, participants indicated that they would have preferred a game with fast loading times and more progression, as they felt the hyper casual game developed was repetitive. This suggests that, in the future, research should be directed towards games that provide users with a sense of progression (i.e., RPG games and adventure games), rather than casual games.

References

[1] A. Wiggins and K. Crowston, "From Conservation to Crowdsourcing: A Typology of Citizen Science," *2011 44th Hawaii International Conference on System Sciences*, 2011, pp. 1-10, doi: 10.1109/HICSS.2011.207.

[2] Alexandra Eveleigh, Charlene Jennett, Ann Blandford, Philip Brohan, and Anna L. Cox. 2014. Designing for dabblers and deterring drop-outs in citizen science. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14). Association for Computing Machinery, New York, NY, USA, 2985–2994. https://doi.org/10.1145/2556288.2557262

[3] Bogost, I. (2011, May 3). *Persuasive Games: Exploitationware*. Game Developer. Retrieved December 13, 2022, from https://www.gamedeveloper.com/design/persuasive-games-exploitationware

[4] Bowser, Anne & Hansen, Derek & He, Yurong & Boston, Carol & Reid, Matthew & Gunnell, Logan & Preece, Jennifer. (2013). Using gamification to inspire new citizen science volunteers. ACM International Conference Proceeding Series. 10.1145/2583008.2583011.

[5] Curtis, V. (2014). Online citizen science games: Opportunities for the biological sciences. *Applied & Translational Genomics*, *3*(4), 90-94. https://doi.org/10.1016/j.atg.2014.07.001

[6] Deterding, S., Dixon, D., Khaled, R., & Nacke, L. (2011). From game design elements to gamefulness: defining "gamification". Paper presented at the proceedings of the 15th international academic MindTrek conference: Envisioning future media environments, Tampere, Finland.

[7] Franzoni, C., & Sauermann, H. (2014). Crowd science: The organization of scientific research in open collaborative projects. *Research Policy*, *43*(1), 1-20. https://doi.org/10.1016/j.respol.2013.07.005

[8] Hamari, J. (2017). Do badges increase user activity? A field experiment on the effects of gamification. *Computers in Human Behavior*, *71*, 469-478. https://doi.org/10.1016/j.chb.2015.03.036

[9] Huang, Biyun & Hew, Khe. (2015). Do points, badges and leaderboard increase learning and activity: A quasi-experiment on the effects of gamification.

[10] Hung, A.C. (2017). A Critique and Defense of Gamification. *Journal of Interactive Online Learning, 15*, 57-72.

[11] Prestopnik, N. R., & Tang, J. (2015). Points, stories, worlds, and diegesis: Comparing player experiences in two citizen science games. *Computers in Human Behavior*, *52*, 492-506. https://doi.org/10.1016/j.chb.2015.05.051