"I want to be a Captain! I want to be a Captain!": Gamification in the *Old Weather* Citizen Science Project

Alexandra Eveleigh[1], Charlene Jennett[1], Stuart Lynn[2], Anna L. Cox[1]
[1] UCL Interaction Centre, University College London, London, UK, WC1E 6BT
[2] Adler Planetarium, 1300 S. Lake Shore Drive, Chicago, IL 60605, USA
alexandra.eveleigh.09, charlene.jennett, anna.cox @ucl.ac.uk
stuart@zooniverse.org

ABSTRACT

Gamification is increasingly implemented in citizen science projects as a means of motivating and sustaining participation. In a survey and subsequent interviews we explored the appeal of gamification for participants in the *Old Weather* project, and its impact upon data quality. We found that the same competitive mechanisms which some volunteers found rewarding and motivating were either ignored by other participants, or contributed to a decision to discontinue participation. We also identified an opportunity to use gamification to exploit the narrative appeal of a project such as *Old Weather*. In contrast to previous citizen science research, much of which focuses on how to support the most active or prolific contributors, we offer new design recommendations which recognise varying levels of engagement with a project.

Author Keywords

Citizen Science; Gamification; Motivation; Engagement.

ACM Classification Keywords

H.1.2. User/Machine Systems: Human Factors.

General Terms

Human Factors.

INTRODUCTION

In citizen science projects, members of the public (citizens) collaborate with professional scientists to conduct scientific research [1]. The tasks that are crowd-sourced to the public are often repetitive. For example, in *Old Weather* [2] volunteers are asked to transcribe page after page of handwritten weather observations. Previous research suggests that gamification - the use of game design elements in non-game contexts [3] - can help to make repetitive tasks more enjoyable [4] and sustain volunteers' engagement [5]. However there are also concerns that gamification could have an adverse effect on data quality [6] and that it might not appeal to all volunteers, some users preferring a 'more serious' interface [7].

In this paper we present our case study of *Old Weather*, which aimed to investigate the appeal of gamification for participants irrespective of the extent of their engagement in the project. First we describe how (and why) gamification was employed in *Old Weather*. Then we describe and discuss our research findings. We conclude with a series of design considerations and challenges for gamifying citizen science projects.

OLD WEATHER

Old Weather [2] was launched by the Zooniverse citizen science consortium in October 2010. Volunteers visit the website where they can read through handwritten pages of 19th century ship log books. Their task is to transcribe the weather observations that were recorded daily. They can also, optionally, transcribe any other information they deem important, such as battle action or movements of personnel. Scientists use the digitized weather data for climate modeling. The digitized logs are used by naval historians for historical research.
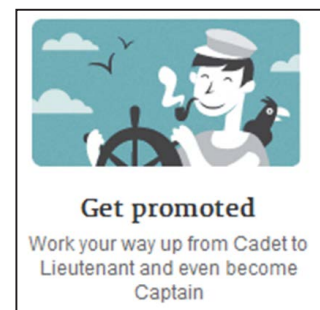

Figure 1. Ranking system used in Old Weather
© Zooniverse 2012

The *Old Weather* team recognized that the volunteers' task was repetitive and the handwriting difficult to read, so they added a ranking system recognizing the level of contribution made by each volunteer as a way of encouraging participation. There are three levels through which a volunteer can progress when they 'join a ship'. Each volunteers starts as a Cadet. Once they have transcribed more than 30 weather observations from one ship they are promoted to Lieutenant. The top transcribers in each ship can then compete to be Captain - the person who has the most contributions for that particular ship. Importantly, this position of Captain can be lost again once

gained (much like the competition to become mayor of a location on *Foursquare*), in contrast to the Lieutenant role which, once achieved, is always retained. The volunteer's personal rank appears on their transcription page. There is also a list of leading 'crew' for each ship.

The Old Weather team decided to gamify in this way to play on the ship theme. They hoped it might encourage loyalty to a particular ship, with volunteers joining and following one ship. This continuity could also help support accurate transcription, as the volunteer would become familiar with the handwriting in each log book.

In February 2013 *Old Weather* was awarded the Royal Meteorological Society's IBM award for 'innovation that matters' [8]. Through the combined efforts of thousands of volunteers, over 1 million ship logbook pages have been transcribed. Our investigation sought to understand the role of the ranking system in motivating volunteers to contribute.

RESEARCH METHODOLOGY

In this paper we draw upon qualitative data from two studies: a survey study (S) that was conducted during July 2012 with 545 respondents; and a semi-structured interview study (I) that was conducted from December 2012 to May 2013 with 18 participants. These studies were conducted as part of our wider work on understanding volunteers' motivations for participating in (and dropping out of) *Old Weather*. An invitation to complete the survey was emailed to all registered users of *Old Weather*, and placed on the project forum. It comprised 16 questions covering participants' background and their motivations for taking part. Follow-up interviews were conducted with selected respondents. In this paper we present only qualitative results relevant to gamification. Themes were derived using the qualitative method Thematic Analysis [9]. To provide further context, we also report the person's total number of transcriptions (t), where available.

RESULTS

Positive Views of Gamification

Validation
In line with past research [5], intrinsic factors (e.g. an interest in naval history and/or climate change) were important motivations for taking part in *Old Weather*. Being part of the project and helping the researchers was rewarding in itself. The ranking system served as an additional reward system, validating volunteers' efforts:

"Well, people like to feel that their contribution has been validated, just as a psychological tool. The same reason that people try to get achievements in video games - it feels like your work has been validated in some way, even if it doesn't have any real world value." (I7 - 3010t)

Tracking Personal Progress
Several interviewees reported using the ranking points *"not*

to brag about it, but just for themselves to see how many classifications or transcriptions they've done." (I1 - 15120t), as milestones to pace themselves against, or to assess their personal contribution towards overall project progress - *"There were statistics for each ship's completion so it felt like I was making more of an impact"* (S - 6070t). Three participants even suggested adding additional stages of progression between Cadet and Captain, to provide an increased number of transcription targets and incentives for individuals to keep contributing:

"I suggested there should be more ranks, it would be more of an incentive. For example, Sub-Lieutenant or top Lieutenant. If there were more targets it would feel like more of an achievement." (I8 - 7123t)

Competition
Interviewees reported feeling compelled to achieve the status of Captain of a ship, and, once this was achieved, to maintain that position:

"I was Captain of a vessel and it felt rather good. Even though it doesn't mean anything per se, to know that I had contributed more log pages than anyone else." (I7 - 3010t)

Unlike other citizen science platforms (e.g. *Foldit, Eyewire*), *Old Weather* does not offer an official table of overall leading contributors, but the participants constructed their own version which had stimulated one interviewee to a degree beyond simply striving for the positions of Lieutenant and Captain:

"There used to be a spreadsheet that somebody had updated […] And I would check that periodically, and it basically ranked contributions. And I would see, oh somebody's got this many more than me; I'm gonna work really hard and then I'll beat them! And oh, I'm really close to them! If I put in another half hour I'll beat them, kinda thing. So I think that helped really, putting in extra effort and getting more data transcribed." (I4 - 17080t)

Negative Views of Gamification

De-motivating
Similar to Massung et al. [11], we found that a simple score-keeping interface can be something of a double-edged sword. High scoring participants are spurred on by vying for the top position, but low scoring participants are simultaneously demotivated by a 'distant competition' they have no hope of reaching:

"I was never even close to it. It seemed like you had to transcribe 10 times as much as I was transcribing, or 100 times what I was transcribing, to even get towards that, so I never gave it much thought. It seemed like it was more of a competition for the harder core users." (I13 - 256t)

With only one Captain for each ship, some people felt that they could never catch up, or were unable to take a break without sacrificing their position, such that they actually felt less motivated to contribute:

"It was kind of a downer to come back and find after a few days that the number of transcriptions necessary to make captain had doubled" (S)

Stressful

There were also negative aspects to being Captain. Trying to retain the status of Captain could sometimes be stressful or exhausting, even for those that admitted that the competition had pushed them forward:

"Yes, I did find it motivated me. I also found it quite stressful. I'm quite a competitive person, and when I got to be Captain of a ship, I wanted to stay there at any cost! And then someone else came along that had more spare time, and so I would get quite stressed trying to stay ahead!" (I8 - 7123t)

Distrust

Two interviewees had suspicions that other volunteers had cheated to remain amongst the top contributors:

"There were people who were… I don't want to say faking or doing wrong classifications, but they were going so fast that you had to say 'oh that can't be correct'. And just to be in the top 10 or number 1." (I1 - 15120t)

Quantity over Quality

Others felt that the ranking system undervalued smaller contributions, or discouraged participants from submitting more detailed transcriptions including daily 'event' occurrences (useful for historical research) in addition to the basic weather observations:

"I think that it is no good that promotion to captain depends on transcribed weather reports. Events are not taken into account; people are interested only in how many weather reports they transcribe." (S - 1564t)

Some volunteers were concerned that the criteria for earning points risked sacrificing quality over quantity:

"If I were in charge of a project like this, I wouldn't promote people to captain or admiral. I would promote them to apprentice seaman, and say, only after you've done so many logs, and only after you've done so many logs reasonably correctly are you now qualified to serve on board." (I5 - 260t)

One participant, noticing that he was transcribing quickly rather than carefully, moved to a less popular ship so that he wouldn't feel the need to compete:

"I noticed as I was checking the progress of the ship, somebody else was logging quite a few hours and logged one or two pages. I realized, just a minute, I'm not here rushing to complete these things as fast as possible. I don't really care who gets HMAS Sydney done. But I felt maybe that my quality would suffer if I was rushing this, you know. […] So then I picked a ship that no-one knew and no-one cared about." (I5 - 260t)

Trivializing the Research Objectives

Not all the interviewees agreed that the competitive aspect of *Old Weather* increased their contributions. Some participants found the race for Captain irritating or trivializing. More occasional project participants generally disregarded the badges, and lapsed participants were sometimes unaware of the points system altogether:

"I was not aware of any scoring in this project, but if there is, then I'm glad I discontinued it. Scoring is for kids…" (S)

Within the research setting of citizen science too, competition may be viewed as running contrary to the serious scientific objectives of the project, and belittling of the volunteers' often considerable efforts (for instance, our interviewees described carrying out extensive online research to establish the location of obscure places mentioned in the logs). One *Old Weather* volunteer made the point that this *"was not an online game, this was a research project"* (I5 - 260t).

Temporal Dynamics

Consistent with Rotman et al. [10], we found that participants' motivations shift over time. With respect to the effectiveness of games-like features specifically, this temporal dynamic may be cyclical:

"I grew a bit hot and cold on it myself. There were certain things where I was determined to hold onto my Captaincy, and sometimes it really annoyed me…" (I3 - 22389t)

The impact may decline as participation moves beyond the initial phase of engagement:

"I will admit to getting a little 'woo hoo!' on my first promotion. But then the difference to the next one was so huge. It was nice to have, but I don't think I would have missed it if it wasn't there." (I15 - 50t)

One possible reason for this may be the same demotivating effect of 'distant competition' [11] which causes some volunteers to drop out of the project altogether:

"I think that in the beginning it's quite fun, but when you don't keep up with your daily contributions [because] you won't have a chance to become Captain." (I16 - 203t)

Narrative Immersion

Another stated source of motivation was following the narrative of the ship chronologically and geographically:

"The 'real' story that those logs imply is as hypnotically fascinating as any form of fiction or non-fiction […] The ship and crew became friends and even the handwriting became clues as to whom had the watch for the day. I ceased actively transcribing partly because […] I was not able to go back to follow the ship on its journeys." (S)

Some participants found themselves hurrying through the transcription task, not necessarily out of a concern to attain promotion to Captain, but just *"to go fast enough to see most of the log pages for that ship"* (S - 7117t). Similarly:

"The rankings don't matter to me and I almost felt they were distracting at first - like a contest. If I did a lot of pages in one go it was just because I didn't want to have missed anything when I logged on again." (S - 1408t)

This shares similarities with narrative immersion (also known as imaginative immersion), which is a key component of gameplay experiences [12]. We suggest there could be an opportunity to take advantage of the narrative appeal of projects like *Old Weather,* to apply game-like features which promote immersion in stories, including the emerging story of the project itself as it progresses.

DISCUSSION AND CONCLUSIONS
We found that the same competitive gamification mechanisms which motivated some leading volunteers were either ignored by more casual participants, or contributed directly to the decision to discontinue participation. This has important implications for offering a balanced range of game-like features in citizen science platforms, so as to support and encourage the most consistently active and prolific contributors, whilst simultaneously seeking to minimize attrition, and to capture the attention of new volunteers trying out the project.

Another key finding is the opportunity to build upon the 'compulsive', 'addictive', 'absorbing', and 'fascinating' aspects of projects such as *Old Weather* that have a narrative appeal. This opens up some interesting possibilities for future research, as gamification in citizen science is currently almost exclusively discussed in quantifiable competition and reward terms, based upon the accrual of points, levels, badges, and scoreboards [6, 7].

We propose the following design considerations for motivating and sustaining participation through gamification in citizen science:

- Ensure any scoring mechanism can provide personal milestone targets alongside competitive incentives. More finely graduated stages of progression would help volunteers feel their contribution is valued at all times.

- Design personalized feedback into the game as a means to recognize quality rather than quantity. Enable participants to assess the accuracy of their contributions, and to learn from and correct their mistakes.

- Recognize that the appeal of game-like features will vary between participants and over time. Interviewees' suggestions for keeping or re-invigorating their interest included scheduled challenges or prize draws, perhaps for teams rather than individual contestants.

- Take advantage of the narrative appeal - *"oh, I can get one more done before I go to bed, alright, I'll do one more"* (I15 - 50t). Allow participants to choose and follow their own narrative path through the project. For *Old Weather,* for instance, this might entail offering transcription choices according to location, or time period, or enabling participants to read through or catch up on other volunteers' transcripts relevant to their personal affiliation to a particular ship.

REFERENCES
1. Newman, G., Wiggins, A., Crall, A., Graham, E., Newman, S. and Crowston, K. The future of citizen science: Emerging technologies and shifting paradigms. *Front Ecol Environment, 10*, 6 (2012), 298-304.

2. http://www.oldweather.org/

3. Deterding, S., Sicart, M., Nacke, L., O'Hara, K. and Dixon, D. Gamification: Using game design elements in non-gaming contexts. *Proc. CHI 2011*, ACM Press (2011), 2425-2428.

4. Flatla, D., Gutwin, C., Nacke, L., Bateman, S. and Mandryk, R. Calibration games: Making calibration tasks enjoyable by adding motivating game elements. *Proc. UIST 2011*, ACM Press (2011).

5. Iacovides, I., Jennett, C. Cornish-Trestrail, C. and Cox, A. L. Do games attract or sustain engagement in citizen science?: A study of volunteers motivations. *Proc. CHI EA '13*, ACM Press (2013), 1101-1106.

6. Bowser, A., Hansen, D. and Preece, J. Gamifying citizen science: Lessons and future directions. *Designing gamification: Creating gameful and playful experiences*, workshop at CHI (2013).

7. Prestopnik, N. & Crowston, K. Purposeful gaming and Socio-computational systems: A citizen science design case. *Proc.Group 2012*, ACM Press (2012).

8. http://blog.oldweather.org/2013/02/20/award-winning/

9. Braun, V., and Clarke, V. Using thematic analysis in psychology. *Qualitative Research in Psychology 3*, 2 (2006), 77-101.

10. Rotman, D., Preece, J., Hammock, J., Procita, K., Hansen, D., Parr, C., Lewis, D. and Jacobs, D. Dynamic changes in motivation in collaborative ecological citizen science projects. *Proc.CSCW 2012*, ACM Press (2012), 217-226.

11. Massung, E., Coyle, D., Cater, K. F., Jay, M. and Preist, C. Using crowdsourcing to support pro-environmental community activism. *Proc. CHI,* ACM Press (2013), 371–380.

12. Ermi, L. and Mayra, F. Fundamental components of the gameplay experience: Analysing immersion. *Proc. DIGRA* (2005)