“Guess what! You’re the first to see this event”: Increasing Contribution to Online Production Communities

Corey Jackson, Kevin Crowston, Gabriel Mugar, Carsten Østerlund
School of Information Studies, Syracuse University
Syracuse, NY 13210
{cjacks04,crowston,gmugar,costerlu}@syr.edu

ABSTRACT
In this paper, we describe the results of an online field experiment examining the impacts of messaging about task novelty on the volume of volunteers’ contribution to an online citizen science project. Encouraging volunteers to provide a little more content as they work should increase the community’s output. Our prior research suggested that an important motivation for participation in online citizen science is the wonder of being the first person to observe an image. To appeal to this motivation, we implemented a prompt during volunteers’ sessions to alert them when they were the first person to see content in an online citizen science project. Our analysis reveals that new volunteers who saw messages increased the volume of annotations contributed to the project. The results of our study suggest an additional strategy to increase work done, particularly by a population who may not return to the project.

CCS Concepts
•Computer systems organization → Embedded systems; Redundancy; Robotics; •Networks → Network reliability;

Keywords
citizen science; online communities; experiment; motivation; intention to treat; novelty

1. INTRODUCTION
Ciampaglia et al. [3] note online voluntary production communities have two options to increase the amount of content generated: increasing the number of new volunteers or increasing the participation and retention of existing volunteers. To remain successful, then managers of online communities need to understand what attracts volunteers to the project and, once volunteers join the project, what factors contribute to the volume of their contribution. The research presented here is an exercise in the latter, increasing the contribution of existing volunteers.

To maximize contribution, online voluntary community managers need to understand what motivates volunteers to join the community and what factors increase volunteers’ motivation to continue contributing once they join the group. Guided by this understanding, they can make motivational aspects of volunteers’ participation more salient and thus increase the likelihood of volunteers contribution. Motivation in online communities has attracted a significant amount of attention in communities like Wikipedia and free/libre open source software (FLOSS) projects. Several studies on Wikipedia, for example, have noted the complex web of motivations for volunteers ranging from recognition of contributions by peers [5], altruism [14].

As well, it is well known that contribution to online communities follows a long-tail distribution (i.e., power law, 80/20 rule) where a handful of volunteers contribute the majority of the content, but most volunteers contribute little. Indeed, most volunteers contribute in only one or two sessions and never return. For instance, in Wikipedia 1% of the editors contribute 55% of edits [7]. One reason for the low contribution volume is the low barrier to participation. In traditional organizations, contributors are bound to the community contractually or receive some other benefits for the contribution while in online production communities, volunteers can join - and leave - with little or no consequence. As a result, many volunteers may try out the community briefly, but then decide it’s not for them. The implication is that encouraging new volunteers to contribute a little more could be particularly productive.

In this paper, we present the results of an online experiment where we showed volunteers to an online citizen science community popup messages that informed them when they viewed data no other volunteer had seen previously. Based on earlier work, we hypothesized that our experimental appeal to participants’ interests would increase motivation and so increase contributions. The research question guiding this experiment is: How does novelty messaging impact volunteers’ contributions to citizen science projects? Though our analysis, we make two important contributions:

- We present the results of an experiment analyzing the effects a stimuli inducing messaging prompt.
- Given the experimental setup, we describe in detail how we analyzed the results of our experiment using an intention to treat analysis, an approach commonly used in medical research, but less seen in GROUP research.

1.1 Citizen Science
Our study is set in the context of an online citizen science project. Citizen science describes projects in which amateur volunteers collect or analyze data to contribute to scientific research. Projects like The Birdhouse Network (TBN) are characterized by a focus of the citizen scientists on collecting or analyzing data. TBN volunteers place nest boxes in their yards and generate and report data they
collected from the nest boxes. Many projects rely on data objects (e.g., images, sound files) submitted by scientists and tasks require volunteers perform tasks on their machines such as filtering or transcribing data objects. The resulting data is then given to scientists to continue the process of investigation. Galaxy Zoo is an example of collaborative projects; scientists provide dumps of imaged galaxies collected by the Sloan Digital Sky Survey telescope and ask amateur volunteers to identify characteristics of the data object which alert scientists to the presence of new galaxies.

One important feature of citizen science projects are the platforms through which amateurs and scientists collaborate. Zooniverse [28] is a citizen science platform which hosts more than forty collaborative projects. The collaboration between professionals and amateurs is technologically mediated through the project website. As with other online voluntary production communities, Zooniverse projects rely on a steady stream of volunteers to complete project tasks.

2. MOTIVATION TO CONTRIBUTE USER-GENERATED CONTENT

How to motivate contributions to online voluntary production communities has received substantial attention from researchers. The link between motivation and contribution is clear and suggests that when volunteers of the community are motivated, a positive impact on content production is observed. Thus, increasing the motivational affordances of an volunteer’s participation is expected to increase the volume of content generated by an volunteer.

The literature on motivation in online communities is quite diverse. Theories such as social loafing, goal-setting, social awareness, social-identity, uses, and gratifications or collaborative effort model have frequently been operationalized in the design of online communities. A major stream of research draws on psychological research on motivations. For example, studies have found that giving volunteers the opportunity to set goals resulted in increased contribution behaviors [10, 33, 1]. Social motives have also been explored. For example, [30] experimented with making workers in Amazon’s Mechanical Turk aware of the presence (social awareness) of other Turkers to increase volunteers’ feelings of attachment (bond-based and identity based) to the group and found when Turkers are assigned to work groups and communication between volunteers of the work groups is supported Turkers show more loyalty to the requester and MTurk.

Studies have examined a range of motives in particular settings. For example, in a survey of open-source programmers Hars and Ou [8] distinguished between external rewards (e.g., peer recognition) and internal rewards (e.g., self-determination) that are the result participation in OSS environments. Lakhani and Wolf [15] found that one of the strongest motivators for OSS developers is how creative they feel when working on the project, as well as intellectual stimulation and the utilitarian motivation of improving programming skills. Oreg and Nov [19] found that reputation-gaining and learning were important motives for contribution to FLOSS. Research has identified somewhat different motivations for contribution to the online encyclopedia Wikipedia. Yang et al. [31] discovered that self-concept motivation was the main factor influencing Wikipedians’ knowledge sharing behaviors. Nov [18] found eight different motivations explaining editing behaviors in Wikipedia, including fun, ideology, career, and social.

Kraut et al. [13] synthesized the research literature to provide a set of design claims that highlight motivational aspects of volunteers’ participation and encourage contribution in online communities. These are grouped in five categories: (1) selection, sorting, highlighting, (2) framing, (3) feedback and rewards, (4) content, task, and activities, and (5) community structure. For example, related to feedback and rewards is the design claim that was receiving sincere feedback about performance increases motivation. By exploring how social science theories and design claims can be incorporated into the design of the community, we might provide useful insight into encouraging volunteers to contribute.

2.1 Motivation in Citizen Science

Turning to citizen science projects more specifically, researchers have identified a variety of reasons volunteers contribute to such projects. In a survey of citizen science volunteers to the Galaxy Zoo project, [21] analyzed more than eighty statements about why volunteers participate in the project. Twelve categories of motivation emerged, including an interest in astronomy (the topic of many of the projects), wanting to help scientists, contributing to science, enjoying beauty, and learning. Reed et al. [23] surveyed 199 users of Zooniverse and identified three factors explaining why volunteers contribute: social engagement with other volunteers, interaction with the site, and helping (or volunteering). The literature on motivation in citizen science projects has also resulted in descriptions of the dynamic nature of motivation suggesting that motivations shift throughout a volunteer’s volunteership [27, 29].

Many of these motivations are similar to other online production communities like Wikipedia and FLOSS. However, citizen science communities are unique in that volunteers annotate images that few people have seen previously. Indeed, given the volume of data, volunteers regularly experience data objects that have not been seen by anyone. For example, the Planet Five shows volunteers images of that have not yet been seen by anyone outside professional astronomy communities. Along with advertising citizen science as a participatory project, many projects also note the possibility of viewing novel images that have not been seen previously. Research suggests that many volunteers are drawn to projects for this reason. Research by Jackson et al. [11] revealed some volunteers are motivated by the possibility to discover new data or find anomalies that others had not previously identified. Additionally, amongst familiar motives like astronomy, contributing to science, the beauty of galaxy images and learning about galaxies, Raddick et al. [21] found volunteers were also motivated by discovery (i.e., “I can look at galaxies that few people have seen before”). Reiss [24] listed sixteen different motives leading to intrinsic feelings. Included in the list was curiosity, described as a desire for knowledge which has potential to lead to wonder.

Being the first to see something does seem to have significant impacts on citizen scientists’ work. One of the most well-known cases is the discovery of a novel object by a Dutch school teacher, which led to the phenomena being named after her: Hanny’s Voorwerp [12]. In Stardust@Home when a user discovers a dust grain, he or she is listed as co-author on the article announcing the discovery and also is given the privilege of naming the dust grain.

These motives have support in research that identifies motives for action such as novelty seeking (neophilia), sensation seeking, and curiosity. Addressing consumerism and the desire for novel products, Campbell [22] writes “There are those neophiliacs whose craving for the new takes the form of a preference for the novel, the strange or even the bizarre. These are the volunteers who appear to place a high value on the stimulus which is provided by the unfamiliar while perceiving the known as boring.” Raymond [22] introduced the idea of neophilia to describe a trait of hackers of being excited and pleased by novelty. Sensation seeking has been researched within the context of internet and technology use. For example, [17] surveyed mobile phone users and found volunteers
who scored high on sensation seeking had a higher frequency of using the phone to make calls and a higher usage of phone features. In crowdsourcing, Law et al. [15] introduced curiosity-inducing stimulus to incentivize Mechanical Turk workers. The researchers showed MTurk workers obscured (i.e., blacked out) visuals that were only revealed when a worker completed parts of the task. Their research showed curiosity improved worker retention.

However, to our knowledge, there has been no research on motivation that manipulates novelty as a way to influence volunteer participation. Based on the prior literature on motivation in citizen science and our past research examining motivation in Zooniverse projects, we believe that highlighting the novelty of participation can lead to increased motivation resulting in increased work. We suspect that for new volunteers mentioning that a data object has not been viewed by other citizen scientists might play to a person’s desire to be first or experience novel occurrences. We therefore offer two hypotheses:

**H1:** Messaging users about the novelty of their experience increases the number of annotations done.

**H2:** Newcomers contribute more in their first session when shown novelty messaging than newcomers contributing in their first session not shown a novelty message.

### 3. SETTING: ZOONIVERSE

The context for this experiment is the online citizen science platform Zooniverse (http://www.zooniverse.org); a web-based platform that hosts more than forty science projects. Volunteers on the site annotate data objects (i.e., images, sound/video recordings, text) to support further scientific research. For example, in Galaxy-Zoo, volunteers record the shape of galaxies; in OldWeather, they transcribe ship logs from World War I. Before analyzing the data objects, volunteers are asked to complete tutorials that explain how to identify relevant characteristics of the data objects. In addition to annotating data objects, many volunteers contribute to other aspects of the community, e.g., visiting blogs or posting comments to forums.

#### 3.1 Project: Higgs Hunters

Our study focuses on the Higgs Hunters project. In Higgs Hunters, volunteers annotate images from the Large Hadron Collider (LHC), a particle collider built to search for the Higgs boson particle. The data objects are images of two beams of particles colliding, creating a shower of new particles, possibly including previously unknown particles, such as the Higgs. Uncharged particles leave a trace in the image; uncharged particles are invisible. Volunteers search the images for decay anomalies or appearances of off-center vertexes, which are indications that a new particle was created but then decayed into other particles (e.g., the Higgs decays only $10^{22}$ seconds after it is created). The annotation interface is shown in Figure 1. Volunteers click on “Off-centre vertex” on the right-hand side, then mark the location of the vertex and how many tracks appeared. Not all data objects have off-center vertexes from particles, so seeing a new particle is akin to finding a needle in a haystack. Each data object is analyzed by multiple volunteers and the different individual annotations merged to find the consensus.

### 4. METHODS

#### 4.1 Experiment Design

To manipulate motivation by appealing to volunteers’ interest in novelty, beginning in October 2014, the Zooniverse introduced a system that alerted volunteers when they viewed a data object that no previous volunteer had seen, specifically, a popup message that read “Guess what! You’re the first to see this event.” This intervention was added for all users, rather than for a randomly selected control and treatment group. As a result, we studied its impact using a quasi-experimental design. Project scientists periodically inject new data objects into the project, resulting in periods where many volunteers see the popup message. We used work done during such periods to create the treatment group. However, eventually every data object has been viewed at least once, and so the popup message is not shown for an extended time as additional annotations are added to the existing data. We used the work done in such periods to create a control group. However, rather than comparing work done by volunteers randomly assigned to treatment or control (a true experiment) we are comparing work done at different times with and without the treatment (a quasi-experiment).

#### 4.2 Data Collection

We collected data from the log files on Zooniverse servers in October 2015. The dataset contained all the annotations done by volunteers up to that time, including a timestamp for each annotation and whether the user was shown the popup message (i.e., if they were the first person to see that data object). The total dataset includes 684,087 annotations contributed by 6,354 volunteers in the Higgs Hunters project, though not all of this data were used in the analysis. Analysis was done at the session level. Annotations were grouped by user and then into sessions, a series of annotations done by a single user that are separated by a gap of less than 30 minutes. The intuition is that a user generally does some number of annotations in a single sitting (possibly with a short break between) and then takes a longer break, e.g., until the next day. We used sessions as the unit of analysis to test the hypothesis that a message might increase interest and so lead to the user’s extending the work done on the system, resulting in a longer session.

For each session, we total the number of data objects annotated and how many popup messages were shown (immediately after new data objects are added to the system it is possible that a user will see many new objects in a single session). Figure 2 shows the number of sessions done per day over time; sessions with at least one popup message are in blue and sessions with no messages in orange. From blue areas in the figure it is easy to identify the dates on which new data objects were added to the system. Note also the very large spike in work done in the first few weeks of the project and the decline in activity over time.
4.3 Within-subjects analysis

We carried out two different analysis on the sessions. Our first analysis was within-subjects, comparing for the same user the length of sessions that had or did not have a popup message. Comparing sessions within a subject help control for the very high variability of contribution to the project. An advantage of the within-subject design is that it uses more of the data. A possible confound to this design is that at certain points in the project, seeing a popup message becomes a matter of chance, as some but not all of the data objects are new. As a result, during those periods, rather than messages causing sessions to be longer, a longer session may have a message simply because “the dice were rolled” more often. To address the skew in the distribution of session lengths, we log transformed the dependent variable. We analyzed data using the nlme package because multilevel models are more sensitive to within-subject variance.

4.4 Between-subjects analysis

To avoid the confound noted above, we also carried out a between-subjects analysis on a subset of the data. To form the treatment group (sessions with at least one popup message), we identified a two-week period during which new data objects were available (the middle light blue bar in the inset in Figure 2). We chose a period later in the life of the project when the number of sessions done per day was beginning to plateau (meaning that the sample is not of early joiners who might be expected to differ from other users) and to avoid the New Years holiday. To eliminate possible influences from prior experiences on the system, we only used sessions from volunteers who had the first session during the period, meaning this analysis is restricted to the impact on newcomers. For this analysis, we included all sessions done by those volunteers in the two weeks after their first session (i.e., the group includes some sessions beyond the treatment period for users who joined late in the period). Note though that not all sessions during this period had a popup message (as shown by the orange part of the graph), an issue we discuss below.

The control group (sessions with no popup message) was formed in a similar fashion. However, because the number of sessions per day was steadily declining over the life of the project, we controlled for maturation by selecting one week just before and one week just after the treatment period. The purple bar in the inset in Figure 2 indicates the first control period, and the last bar, the second control period. Because we followed volunteers for two weeks after their first session, we left a week gap between the first control period and the treatment period. Even so, we had to drop two sessions from the control group that edged into the treatment period and included a popup message. Both control and treatment periods were multiple of 7 days to include all days of the week equally.

5. RESULTS

The contributions of volunteers in citizen science projects typically follow a long-tail distribution: many volunteers contribute little content, e.g., only a single session while a dedicated handful of volunteers contribute the majority of content. Higgs Hunters volunteers are no different: most volunteers (71%) contribute in only one session. The average number of sessions by volunteers is 2.73. Since the number of sessions is likely to be skewed we computed a .1 percent trimmed mean to illustrate how little work most volunteers contribute and discovered the average session dropped to 1.32 (std. dev. = 11.52) and the average number of annotations dropped significantly from 107.65 annotations to 30.47 (std. dev. = 699.36). Conversely, eight volunteers have contributed more than 11,000 annotations each - a stark contrast in behavior.

5.1 H1:Popups Increase Contribution

5.1.1 Within-subjects analysis

We analyzed the effects of a message on the number of annotations submitted during sessions. In total, N volunteers had both types of sessions (messages seen and no messages seen). The number of sessions where messages were seen was 1,355 while 5,618 sessions had no messages. The message sessions had an average of
missions on the participant, the message has a significant impact on the measured annotation sub-
tions (std. dev. = 77.8) and had a trimmed mean of 25.5. Seeing a 77.5 (std. dev. = 107.6) and a trimmed mean of 55.4 annotations population, ** indicates computed values based on (observed PC and assumed PC-NT, and *** assumed to be same as T-NT.

Table 2: Descriptive statistics for treatment and hypothetically derived control groups. In the table * indicates observed values from T-T group is not random, but is related to the outcome variable, is more likely to have a popup. As a result, volunteership in the chance, so a longer session (i.e., from a more interested volunteer)

<table>
<thead>
<tr>
<th>Experiment Group</th>
<th>Start-End Date (No. days)</th>
<th>No. Users</th>
<th>No. Sessions</th>
<th>Average No. Annotations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control Prior (CPrG)</td>
<td>1/8/15-1/14/15 (7 days)</td>
<td>107</td>
<td>144</td>
<td>18.1 (std. dev. = 26.1)</td>
</tr>
<tr>
<td>Control Post (CPG)</td>
<td>2/5/15-2/11/15 (7 days)</td>
<td>76</td>
<td>141</td>
<td>36.8 (std. dev. = 53.6)</td>
</tr>
<tr>
<td>Pooled Control (PCG)</td>
<td>14 days</td>
<td>183</td>
<td>285</td>
<td>27.3 (std. dev. = 43)</td>
</tr>
<tr>
<td>Treatment (TG)</td>
<td>1/21/15-2/3/15 (14 days)</td>
<td>217</td>
<td>356</td>
<td>46.4 (std. dev. = 74.7)</td>
</tr>
</tbody>
</table>

Table 1: Descriptive statistics of treatment and control groups.

<table>
<thead>
<tr>
<th>Experiment Group</th>
<th>No. Sessions</th>
<th>Average No. Annotations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment (T)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treated (TT)</td>
<td>223 (62.6%)*</td>
<td>64.7 (std. dev. = 84.6)*</td>
</tr>
<tr>
<td>Not treated (T-NT)</td>
<td>133 (37.4%)*</td>
<td>15.7 (std. dev. = 37.9)*</td>
</tr>
<tr>
<td>Pooled Control (PC)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hypothetically treated (C-T)</td>
<td>179 (62.6% of 285)**</td>
<td>34.4 (std. dev. = 44.4)**</td>
</tr>
<tr>
<td>Hypothetically Not treated (C-NT)</td>
<td>106 (37.4% of 285)**</td>
<td>15.7 (std. dev. = 37.9)**</td>
</tr>
</tbody>
</table>

77.5 (std. dev. = 107.6) and a trimmed mean of 55.4 annotations while the no messages sessions submitted on average 41.6 annotations (std. dev. = 77.8) and had a trimmed mean of 25.5. Seeing a message has a significant impact on the measured annotation submissions on the participant, $\chi^2(2) = 171.9, p < 0.0001$.

5.1.2 Between-subjects analysis

In the between-subject analysis, we compare the length of sessions done by those who joined during a treatment period with many popup messages (the blue bar in Figure to those who joined during a control period without such messages (the purple bars). During the treatment period, 217 new volunteers joined and contributed across 356 sessions (TG). During the first control period, 107 new volunteers joined the project and contributed annotation in 144 sessions and the second period, 76 new volunteers joined and contributed in 141 sessions, for a combined total of 183 new volunteers contributing in 285 sessions during the pooled control periods (PCG). The average number of annotations done in those sessions is shown in the final column of Table 1. A t-test shows that the difference between the average sessions length in the TG (46.4) and PCG (27.3) is statistically significant ($t(639) = 3.83, p < .001$), suggesting that the popup messages do increase the length of a session.

The estimate above is conservative, since as noted above, not all session in the treatment group actually experienced the treatment. Specifically, of the 356 sessions in TG, only 223 had a popup message (62.6%); the remaining 133 (37.4%) did not. As a result, our estimate of the impact of the treatment is diluted by the sessions in which the treatment was not seen. The 223 sessions in the subset of sessions that were treated (TT) had 64.7 (std. dev. = 84.6) annotations on average, while volunteers’ sessions in the treatment but not treated (T-NT) subset contributed only an average of 15.7 (std. dev. = 37.94) annotations.

To get a better estimate of the impact of a popup message, we apply an intention to treat analysis. Figure 3 shows how the subjects in the analysis are divided into control and treatment group and how the treatment group is further divided into treated and untreated groups. We can observe the average length of a treated session (T-T, 64.7 annotations), so the problem is to find a suitable comparison group of untreated sessions. We can not simply compare T-T to T-NT or the entire control group (PCG) because of the confound noted above: at some times seeing a popup message is a matter of chance, so a longer session (i.e., from a more interested volunteer) is more likely to have a popup. As a result, volunteership in the T-T group is not random, but is related to the outcome variable, meaning the difference between T-T and the other groups could be due to selection rather than the treatment (as shown in Figure 3).

Figure 3: Logic of intention to treat analysis. Regular font indicates observed groups; italics indicate hypothesized groups.

To create a comparison group for T-T we need to select a comparable subset of the control group. Fortunately, it is not necessary to actually carry out the selection; instead, we can do it hypotheti-
cally and compute the results. We assume that the control and treatment groups are identical aside from the treatment. Such comparability is the goal of experimental design and an assumption of a quasi-experimental design. Therefore, had the control group been treated, it would have split in the same proportion as T into a subset that would have received the treatment (C-T, e.g., sessions from more interested control group volunteers) and a subset that would not have received the treatment (C-NT, e.g., sessions from less interested volunteers). As the volunteers of the C-T subgroup are selected in the same way as the T-T subgroup, they should be comparable to the T-T-subgroup, aside from the treatment; and similarly for the C-NT and T-NT subgroups.

We compute the properties of the C-T subgroup indirectly. Since they are identically selected subsets of assumed-to-be identical groups, T-T and C-NT are assumed to have identical properties: the same mean number of annotations (15.7) and same standard deviation (37.9). Given the observed properties of C as a whole and the assumed properties of the C-NT subset (the same as T-NT), we can compute the properties of the C-T subgroup to compare to T-T. The results are shown in Table 2. A t-test comparing the mean number of annotations in T-T (64.7, std. dev. = 84.7) to the hypothesized number in C-T (34.3, std. dev. = 44.4) shows a statistically significant difference, $t(400) = 4.35, p < .001$. As expected, this differ-
Our analysis confirmed both our hypothesis. First that novelty messaging increases contributions. We showed that messages alerting volunteers when they saw novel data objects increased the number of annotations submitted by volunteers from 34.4 to 64.7 in a session - almost twice as many annotations were submitted by the treatment groups. We also showed an ability to extend sessions for the same volunteers (within-subjects) and showed in session where volunteer saw a message they contributed significantly more annotations (77.5 vs. 41.6). Second, we showed that the population most likely to leave can be motivated to contribute more prior to their departure. In volunteers’ first sessions our experimental manipulation increase from annotations from 24.9 to 51.9 - doubling their departure. In volunteers’ first sessions our experimental manipulation had an impact on a population who was unlikely to contribute in future sessions. Such a comparison is interesting because as noted earlier, many volunteers contribute only one session, so increasing the length of this session may have a big impact on the project. This comparison included 135 sessions in the treatment treated subset (TTN-Newcomer), 82 in the treatment not-treated (T-NT-Newcomer), 114 in the hypothesized pooled control who would have been treated (HPC-T-Newcomer), and 69 in the hypothesized pooled control which would not have been treated (HPC-NT-Newcomer). The descriptive statistics for the population of initial sessions are shown in Table 3. We performed a similar intention to treat analysis on the population of newcomers. Treated newcomers (TTN-Newcomer) contributed 51.9 annotations (std. dev. = 69.3) while those in the hypothesized control (those who would have been treated, HPC-T-Newcomer) contributed only 24.9 (std. dev = 34.4) annotations, a statistically significant difference of 27 annotations, t(247) = 3.8, p < .001.

Table 3: Newcomer experiment groups with outcome variable annotations. In the table * indicates observed values from population, ** indicates computed values based on hypothesized control group, and *** assumed to be same as T-NT.

<table>
<thead>
<tr>
<th>Experiment Group</th>
<th>No. Sessions</th>
<th>Average No. Annotations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment Newcomer (TN)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treated (TTN-Newcomer)</td>
<td>135</td>
<td>51.9 (std. dev. = 69.3)</td>
</tr>
<tr>
<td>Not-treated (T-NT-Newcomer)</td>
<td>82</td>
<td>7.1 (std. dev = 10.9)</td>
</tr>
<tr>
<td>Hypothesized - Treated (HPC-T-Newcomer)</td>
<td>114</td>
<td>24.9 (std. dev = 34.4)</td>
</tr>
<tr>
<td>Hypothesized - Not treated (HPC-NT-Newcomer)</td>
<td>69</td>
<td>7.1 (std. dev = 10.9)</td>
</tr>
</tbody>
</table>

6. DISCUSSION

Our analysis confirmed both our hypothesis. First that novelty messaging increases contributions. We showed that messages alerting volunteers when they saw novel data objects increased the number of annotations submitted by volunteers from 34.4 to 64.7 in a session - almost twice as many annotations were submitted by the treatment groups. We also showed an ability to extend sessions for the same volunteers (within-subjects) and showed in session where volunteer saw a message they contributed significantly more annotations (77.5 vs. 41.6). Second, we showed that the population most likely to leave can be motivated to contribute more prior to their departure. In volunteers’ first sessions our experimental manipulation increase from annotations from 24.9 to 51.9 - doubling the number of annotations.

6.1 Implications for Design

The literature on motivation in online production communities like the one researched in this paper point to a variety of strategies to motivate volunteers to contribute to projects. For the volunteers in citizen science projects, novel data objects have real implications. For example, in another Zooniverse project - Seafloor Explorer volunteers coalesce around a “stripey tube-dwelling creature” after volunteers asked, “what are the tube shape and the stripey creature?” Once individuals surrounded the data object to respond to the citizen scientists, it was discovered that a new species - named convict worm1 by the citizen scientists was discovered in the dataset. Across other Zooniverse projects, similar stories are present. Our results have clear implications for system designers, specifically that they might want to: a) include such a message if possible and b) spread novel experiences across users and sessions to maximize their impact, rather than having one lucky user see many of them. We also found that for a population that, statistically, is not expected to return, we were able to increase the number of annotations submitted, proof that our messaging makes the project more appealing.

Like previous studies we provide evidence that motivational artifacts can be designed and implemented in online communities to encourage volunteers to contribute. While the most apparent uses of our popup might be in citizen science where the ecosystem supports motivation through novel data objects, we can also imagine how novelty might appear in other communities such as Wikipedia, open source software projects, Q&A communities, or blogging sites. In Q&A communities, being the first to respond to a post hold promises of increased attention to one’s comment with the hopes that others after the initial poster will see their responses. In fact, a study of Answerbag, a Q&A site, Gazan et al. found the first submitted answers accumulate 17 percent more rating points than subsequent responses. Gazan et al. also noted: “If there is a first-mover advantage in a social Q&A environment, there must be a measurable benefit to having the equivalent of a dominant market position, regarding some desirable limited good.” The first mover sets the topic of conversation or is perhaps the first individual to point out a novel feature (i.e., in citizen science). Another example is presented in Wordpress pages where authors are encouraged to publish posts with a “Be the first to comment” script at the end to encourage readers to start a conversation. In online communities where answers are valued highly being the first to post might make a comment more prominent to readers or in communities where social voting is a feature, increase the number of up-votes in the community. In Wikipedia, community builders might encourage new editors to be the first to create a new article. In open source software communities designers of the projects might highlight novel coding challenges which other members have yet to see to make one the first to solve the problem. There is some existing linkage to being the first and motivation (e.g., building social capital). From our reading of literature, one might call our experiment novelty seeking, discovery, or curiosity. To our knowledge, no research on has addressed making novel experiences known to users as a method to increase motivation and, in turn, increase contribution. More research is needed to conclude what this phenomenon is - novelty seeking, discovery, curiosity, sensation seeking, neophilia, some other term, or maybe some combination of existing names.

1http://blog.seafloorexplorer.org/tag/convict-worm/
6.2 Encouragement Designs for Similar Experiments

One of the main contributions of this research was the use of an intention to treat approach to analyzing the data. More about the approach can be found here: [9]. The approach was necessary because not all of the sessions selected for the treatment group actually received the treatment. We believe that other interventions in the working of online communities may face the same limitation, meaning that this approach to analyzing the data may be generally useful.

7. LIMITATIONS

There are four limitations we wish to alert the reader to. First, is the experiment design itself: a quasi-experiment. Given the irregular introduction of the treatment to the population, we had to find an alternative approach to analyzing the data the was generated from the experiment. A true randomized controlled experiment would have been preferred, but we faced limits on our ability to manipulate the site in the way that would be required to implement a randomized control. Researchers studying other online communities might also have limited capabilities.

Second, there is the possibility that the treatment lasts longer than the sessions we analyzed. For example, while a message could be viewed during a volunteer’s first session, the effects of that message might have motivational impacts on the next session since the volunteer might have returned because of the message in the first session. Third, and perhaps specific to citizen science projects, if the annotation task takes time to learn (some projects are harder than others), it maybe that the work contributed by the newcomers isn’t that valuable and lengthening their first session isn’t so useful to the project. Nonetheless, most projects here rely on the wisdom of crowds [29] so incorrect annotations are likely to be crowded out by correct annotations.

Lastly, we did not determine whether seeing multiple messages in a session increases the impact of our treatment. We consider this a limitation, but also future work. Given the discussion of curiosity, novelty, discovery it will be interesting to determine to what extent the messages have a diminishing effect on motivation or possibly decrease motivation. Lastly, as with most between-subjects designs, our analysis failed to account for differences in volunteers. In crowdsourcing projects, where low barriers to exit exist, it is unknown whether volunteers who left the project did so simply because a lack of experiencing the treatment.

8. ACKNOWLEDGEMENTS

Without the Zooniverse volunteers working on the projects and taking the time to answer our questions there would be no paper. Many thanks to the Zooniverse team implementing the experiment. This material is based upon work supported by the National Science Foundation under Grant No. xx-xxxxx.

9. REFERENCES


