Which motives are most effective in recruiting citizen scientists? Results of a field experiment

Tae Kyoung Lee[1], Kevin Crowston[2], Mahboobeh Harandi[2], Carsten Østerlund[2], Grant Miller[3]

[1] Department of Communication, University of Utah [2] School of Information Studies, Syracuse University [3] Zooniverse, University of Oxford

ABSTRACT: *Researchers have identified a range of motives for participation in citizen science projects, but have not investigated their relative efficacy in encouraging initial participation. We report on an experiment (N=36,513) that compared the response to emails designed to appeal to four different motives for participation. We found that the messages appealing to the possibility of contributing to science and learning about science attracted more attention than did one about helping scientists but that one asking for help for scientists generated more initial contributions. Overall, contributing to science resulted in the largest volume of contributions and joining a community the lowest.*

Keywords: citizen science.

Introduction

Citizen science refers to scientific projects that receive voluntary contributions from members of the general population. Depending on the project, contributions range from collected data (e.g., bird observations in the eBird project) to analyses or annotations of already-collected data (e.g., in Zooniverse projects such as the Gravity Spy project examined in this paper). In any case, the success of citizen science projects is heavily dependent on attracting participation from citizen scientists. Many studies have conducted surveys and interviews with citizen scientists to identify the motivations for their participation [e.g., Curtis, 2015; Land-Zandstra, Devilee, Snik, Buurmeijer, & Broek, 2016a; Nov, Arazy, & Anderson, 2011; Raddick et al., 2010; Wright, Underhill, Keene, & Knight, 2015]. Such research suggests factors that are important in motivating citizen scientists to contribute. However, few studies, if any, have examined how to appeal those motivations in recruiting volunteers.

To attract citizen scientists to participate in a project, researchers, first, need to let them know about it, what Scheliga, Friesike, Puschmann, and Fecher [2016] called "crowd building". In many cases, researchers send an invitation email to potential citizen scientists or present a message on a citizen science project platform. Although recruiting messages are the first point of contact with new participants and research suggests that certain motivations are reported more often among participants, researchers do not seem to have identified which messages are more effective in attracting participants. Studies by Robson, Hearst, Kau, and Pierce [2013] and Crall et al. [2017] have examined the efficacy of different media in recruiting participants, e.g., by comparing traditional media and social networking, but not the efficacy of different messages. Further, given that motivations for participation change during a volunteer's engagement with a project, it is not clear what kind of initial appeal will make a difference to a volunteer's involvement over time.

To fill the gap, we first review literature on citizen scientists' motivation and identify four motivations that might be appealed to when contacting potential citizen scientists in a recruiting message. Then, we create messages to appeal each of the motivations and test how each motivation is associated with participation at different stages in a citizen science project.

Theory: Motivations of Citizen Science Volunteers

Based on recent research on citizen scientists' motivations [e.g., Curtis, 2015; Kaufman, Flanagan, & Punjasthitkul, 2016; Raddick et al., 2010; Reed, Raddick, Lardner, & Carney, 2013; Rotman et al., 2012], we identified four motivations that have been found to be important to citizen science volunteers: learning science, joining a community, contribution to science, and altruism. We chose these four because they were frequently reported as being motivators but with different findings about their effectiveness, so a study comparing their appeal is needed. On the other hand, we had to exclude motivations that could not be manipulated in the experiment. For example, even though many volunteers report being motivated by an interest in the science of the project, we could not test differences in this motivation in our study, since we had only one project on one topic.

Learning about science

Citizen science projects range various science topics, from biology to astronomy. To help citizen scientists understand a project, researchers typically provide detailed information about the area of the project. Many projects are explicitly designed to have citizen scientists experience the scientific processes. Bonney et al. [2009] concluded that most citizen science projects are designed to help citizen scientists learn scientific knowledge to some degree. In other studies, citizen scientists learned relevant science knowledge while participating in the project.

Consistent with these efforts, studies have identified learning science as one of the major motivations for volunteers in participation [e.g., Brossard, Lewenstein, & Bonney, 2005; Land-Zandstra et al., 2016a; Masters et al., 2016]. Kraut and Resnick [2011] similarly argued that since citizen scientists are not provided with monetary rewards, getting knowledge about science can be rewards to encourage them to participate, so whether people can learn science or not by participating in a project is important. Rotman et al. [2012] state that volunteers report in interviews being motivated by "the opportunity to learn more and widen their scientific horizons". Cox et al. [2017] found that understanding motives, which include learning, were associated with more contributions. Domroese and Johnson [2017] found learning about bees to be the most cited reason for participating in the Great Pollinator Project.

Contributing to science

Citizen science projects are designed to contribute to the scientific process. Recent surveys and interviews show that citizen science volunteers are motivated to participate in projects by the opportunity to contribute to science [e.g., Brossard et al., 2005; Land-Zandstra et al., 2016a; Land-Zandstra, van Beusekom, Koppeschaar, & van den Broek, 2016b; Reed et al., 2013]. The possibility to contribute to science has emerged as a major motivation in several studies. For example, Zooniverse volunteers answered that they are more motivated by their contribution to science than by the possibility to learn about science or altruism [Brossard et al., 2005; Reed et al., 2013]. Contributing to science was listed as a primary contributor for participation in CosmoQuest [Gugliucci, Gay, & Bracey, 2014], Foldit [Curtis, 2015], and in the Dutch Great Influenza Survey [Land-Zandstra et al., 2016b] and the second most cited reason for the Great Pollinator Project [Domroese & Johnson, 2017]. More interestingly, Land-Zandstra et al. [2016b] found that citizen scientists who had participated in the project for a longer time were more motivated by contributing to science.

Joining a Community

As social creatures, humans seek the community of others. Accordingly, researchers suggest that citizen scientists are sometimes motivated to engage in projects to join a community. Conversely, when people notice that many other people engage in some activity, they perceive it as a social norm to follow [Kraut & Resnick, 2011] and are therefore motivated to engage in the same activity, a phenomenon called social proof [Cialdini, 2001] or social norm [Kaufman et al., 2016]. For both reasons, volunteers may be motivated to join a project that they know others are part of.

Evidence for the effect of community or social proof on citizen scientists' participation is inconsistent. Holohan and Garg [2005] studied distributed computing projects, which require a very minimal commitment, and found that while only a small fraction of contributors were members of teams, the team members were among the largest contributors, which they took as evidence for the power of community. In an interview study of participants in FoldIt, citizen scientists' desire to be a part of the community emerged as a motivation to participate [Curtis, 2015]. On the other hand, Rotman et al. [2012] report that "community involvement was not mentioned as a primary motivation for participation in scientific projects". Cox et al. [2017] actually found a negative relation between social motives and volume of contribution. In a recent experiment study, Kaufman et al. [2016] showed that appealing to social proof was less effective than appealing to altruism in encouraging people to participate in a project. They argued that when people see many people already participating in a project, they do not make much efforts due to social loafing.

Altruism

Finally, helping scientists is suggested as a motivation for citizen scientists. Citizen science projects are designed by professional scientists and rely contributions to citizen scientists to achieve the goal of the project. Thus, making contribution to the projects is an altruistic act that helps professional scientists achieve their goals. Crowston and Fagnot [2008] adopted a model of helping behaviors to explain motivations underlying massive virtual collaborations.

Altruism was found effective in leading people to contribute to a crowdsourcing game [Kaufman et al., 2016] and in an open source project, altruism was found as participants' prominent motivation [Oreg & Nov, 2008].

Motivations at different stages of contribution

An important consideration in studying how different motivations appeal to volunteers is that effective motivations change as participants get to know the project. Crowston and Fagnot [2008] argued specifically that the motivation for initial contribution to a collective project are different than the motives for sustained participation, a finding echoed by Rotman et al. [2012]. Accordingly, we consider that motives might differ for the decision to participate, initial participation and sustained participation. However, different studies make different suggestions about which motives are salient at different stages. For example, Cox et al. [2017] found that an "understanding motivation [i.e., learning] associates even more strongly and positively with volunteering at higher percentiles of activity" (that is, for volunteers who have contributed more), while West and Pateman [2016] report that "social factors were significant in retaining volunteers in the long-term", and further, that initial motives matter, as "people with certain motivations more likely to continue volunteering than others".

Present Study

In summary, prior research has identified a range of motives for contribution to citizen science projects. However, it is difficult to draw a clear picture of the relative effectiveness of different motives from this work. As a result, it is not clear which motivations should be appealed in a message to recruit participants. Further, the evolution of a volunteer's participation and motives means that results may depend on when they are measured. Thus, we test the relative efficacy of messages appealing to each motivation to answer the following research question: Among messages appealing to four motivations that been identified as important to citizen science, which is the most effective? Specifically, 1) which message attracts the highest number of volunteers and 2) which attracts the highest number of contributions from volunteers?

Methods

Setting: The Zooniverse Citizen Science Platform

Our empirical study is set in the context of an online citizen science project. While there are several models of citizen science, the project we investigate here involves volunteers in large-scale scientific data analysis. Such citizen science projects rely on an online worldwide collaboration platform to support the involvement of scientists and the public. The scientists share their research projects with the public who are interested in the science.

More specifically, we draw on data from Zooniverse. Zooniverse is the largest platform for citizen science projects, hosting more than 70 individual projects at the time of writing, in astronomy, history, oceanography, and many other fields. In Zooniverse projects, scientists upload data objects to the platform and ask a series of questions to collect information about the objects or help filter useful data objects from those which might not be useful for the scientists.

The project we studied is Gravity Spy [Zevin et al., 2017]. Zooniverse, the Laser Interferometer Gravitational-Wave Observatory (LIGO) Scientific Collaboration (LSC) and citizen science researchers launched this project in October 2016. The goal of Gravity Spy is improving the interferometers used to search for gravitational waves. A challenge for LIGO scientists is that the detectors need extremely high sensitivity to be able to detect gravitational waves, but as a result, the detectors also record a large quantity of noise (referred to as glitches). The glitches obscure or even masquerade as gravitational wave signals, reducing the efficacy of the search. Currently there are more than 20 known classes of glitch with different causes, with the possibility of more classes being identified as the detector is worked on. Gravity Spy recruits volunteers to classify glitches into the known or novel classes. Having a collection of glitches of the same class helps to focus the LIGO scientists' search for their source.

Study Design and Procedure

Zooniverse project staff routinely email members of a mailing list to announce new projects and to solicit contributions. For this experiment, we created four versions of an email message recruiting new volunteers for the Gravity Spy project. These messages were the first public announcement of the project to the list; it had earlier been in beta test with a more select group of participants. The project was simultaneously announced via other channels, attracting new volunteers who did not receive one of these messages. All four messages provided the same short description of the new project but differed in the first and last sentences, which were tailored to emphasize one of the motives discussed above. The first sentences of each message were as follows (not including the phrase in bold italics):

1. ***Science learning***: Extend your knowledge in astrophysics by participating in Gravity Spy!
2. ***Community***: Join your fellow citizen scientists in classifying problematic noise in the search for gravitational waves!
3. ***Contribution to science***: You can contribute to science by classifying problematic noise in the search for gravitational waves!
4. ***Altruism***: Astrophysicists need your help to classify problematic noise in the search for gravitational waves!

The full text of each message is included in the appendix. As with other Zooniverse announcement emails, included in the message was a unique link to the Gravity Spy home page for each individual recipient, which allowed the Zooniverse staff to track if a message recipient visited the website by clicking on the link provided.

For the experiment, mailing list members were randomly assigned to one of four cohorts (one per message). The cohorts had between 9,123 and 9,131 members, as shown in Table 1 (below) for a total of 36,513 recipients. The numbers in the cohorts differed due to changes in the mailing list during the experiment. The assigned recruiting emails were sent to users on the Zooniverse email list on 12 October 2016. The process of sending emails takes several hours, so different users receive the email at different times during the day and of course, we cannot be certain when the message was read.

Data

Three weeks after the messages were delivered, we collected the number of clicks on the links in the emails to the project site and the number. On 31 January 2017, we collected the classifications done on the Gravity Spy system by all volunteers who had joined the project after the messages were sent. Data for the users were divided into five groups: one for each of the cohorts who had been sent a recruiting message and a fifth group for new volunteers who had joined during that time but who had not been sent a message (i.e., those not on the mailing list).

Ethics review

The plan for our experiments was reviewed by Syracuse's IRB. A section of the initial volunteer agreement when volunteers sign up for Zooniverse is disclosure that site administrators run experiments to improve the system and volunteer experience. Zooniverse members opt-in to being on the mailing list. The recruitment process was the same as for any Zooniverse project aside from the minor changes in wording. The procedure posed minimal or no risk to the participants. The study does not use any information about the volunteer aside from their behaviours on the site. The site does not collect demographic information of any kind and volunteers are identified only by a self-selected volunteer ID. Collecting informed consent for the experiment would be practically infeasible, given the nature of the study, which is based on emailing members of the mailing list. We were therefore permitted to run the experiment without collecting specific informed consent for participation.

Results

Table 1 shows data about the response to the emailed recruiting messages for the four cohorts. In addition, 2,808 volunteers who did not receive a recruiting message joined and contributed during the experimental period.

Question 1: Which message attracted the highest number of volunteers?

We answer question 1 in three ways, corresponding to the three stages in a new volunteer's movement into participating in the project: decision to participate, initial participation and sustained participation.

Decision to participate

First, as noted above, each message sent included a unique link to the project that enabled the Zooniverse team to track responses. We counted how many of those links had been clicked (shown as Click through in Table 1), indicating that the volunteer decided to visit the project because of the message. We cut off data collection three weeks after the message was sent, as the growth in the number of clicks had ended at that point.

Cohort	Messages sent	Click throughs		Contributors	
		Count	Percent	Count	Percent
Altruism	9,131	429	4.70%	223	56.7%
Contribution to Science	9,129	508	5.56%	215	48.4%
Community	9,130	490	5.37%	176	38.3%
Science Learning	9,123	529	5.80%	194	38.2%
Total of 4 cohorts	36,513	1,956	5.35%	808	45.4%

Table 1. Response statistics for the 4 cohorts who received messages. Contributors are those who made a classification on the site. Contributors percentage is the count of contributors divided by "click throughs".

To determine which messages attracted more volunteers to visit the site, we performed a differences of proportion test comparing the click-through percent for each pair of messages. The z-score and p values for each comparison are shown in Table 2. Because we ran multiple tests, we applied a Bonferroni correction to the significance of each test. According to Sidak's adjustment, to maintain an overall alpha of 0.05 for the collection of 6 tests, each individual test should have an alpha of 0.0085. With the correction, the difference of proportion tests shows that messages *Science Learning* and *Contribution to Science* attracted significantly more click through than *Altruism*, while the other differences are not significant. The final column shows the 99.15% confidence interval for the difference (i.e., with the same correction for multiple tests). The range of the intervals are smaller than 2%, suggesting that the lack of significant results reflect a small difference rather than a lack of power in our tests.

Message Pairs	Difference	$\chi 2$	Unadjusted p-value	Confidence Interval (99.15%)
Science Learning—Contribution to Science	0.23%	0.423	0.515	-0.67%, 1.14%
Science Learning—Community	0.43%	1.270	0.216	-0.47%, 1.33%
Science Learning—Altruism	1.10%	10.889	0.001*	0.22%, 1.97%
Contribution to Science—Community	0.20%	0.308	0.579	-0.69%, 1.09%
Contribution to Science—Altruism	0.87%	6.862	0.008*	0%, 1.73%
Community—Altruism	0.67%	4.180	0.04	-0.18%, 1.53%

Table 2. Results of tests comparing the proportion of message recipients who clicked on the link to the project between each pair of message conditions. * Difference is significant at p<0.05 after Bonferroni correction.

Initial participation

The above analysis examined how many users visited the site after receiving a message. However, only a fraction of those who visited the site went on to actually contribute to the project by making classifications. The number of the message recipients in each cohort who did a classification is shown in the Contributors count column of Table 1. We ran the same proportion test comparing cohorts on the fraction of the visitors who became contributors (the Contributors percentage column, computed as the number who contributed divided by the number who clicked through from Table 1) with the same Bonferroni correction. The results are shown in Table 3.

Message Pairs	Difference	χ^2	Unadjusted p-value	Confidence Interval (99.15%)
Contribution to Science—Science Learning	6%	3.231	0.072	-2.5%, 13.8%
Science Learning—Community	1%	0.034	0.853	-7.3%, 8.8%
Altruism—Science Learning	15%	21.967	0.000*	6.7%, 23.9%
Contribution to Science—Community	6%	4.029	0.044	-1.9%, 14.7%
Altruism—Contribution to Science	10%	8.332	0.003*	0%, 18.4%
Altruism—Community	16%	23.375	0.000*	7.3%, 24.8%

Table 3. Results of tests comparing the proportion of visitors who made contribution between each pair of message conditions. * Difference is significant at $p<0.05$ after Bonferroni correction.

The results show that there is a statistically significant difference in the fraction of visitors to the site who go on to contribute to the project. The percentage for *Altruism* is higher than for all three other cohorts, but the other differences are not significant. Specifically, even though the message appealing to altruism had the lowest proportion of click-throughs, a significantly higher fraction of the volunteers who clicked on the link in that message went on to contribute to the project.

Sustained participation

Finally, we considered how many volunteers became sustained contributors. For this analysis, we aggregated each volunteer's classifications into sessions, defined as a sequential set of classifications separated by a gap of not more than 30 minutes [Mao, Kamar, & Horvitz, 2013]. The intuition is that volunteers tend to come to the system, do one or more classification in a short period with a short gap between classifications, then take a break until later (e.g., the next day), leaving a longer gap between the classifications, which defines a session boundary. The summary statistics for the session analysis are shown in Table 4.

An indication of sustained contribution is a larger number of sessions. We also show the number and fraction of volunteers who contributed more than one session (computed as count of volunteers with more than 1 session divided by the number of contributors from Table 1). Given that most volunteers contribute to a project just once (note that the median number of sessions in all cohorts is 1, i.e., "one and done", [McInnis, Murnane, Epstein, Cosley, & Leshed, 2016]), another indication of sustained contribution is whether the volunteer comes back for a second session.

Cohort	**Sessions**				**> 1 session**	
	Total	**Mean**	**Median**	**SD**	**Count**	**Percent**
Altruism	854	3.8	1	8.3	98	43.9%
Contribution to Science	1268	5.9	1	16.4	97	45.1%
Community	691	3.9	1	9.2	73	41.5%
Science Learning	827	4.3	1	13.1	70	36.1%
Total of 4 cohorts	3640	4.5	1	12.3	338	41.8%
Non-cohort	8883	3.2	1	10.6	966	34.4%

Table 4. Contribution statistics for experimental groups: number of sessions for volunteers in the 4 cohorts who received messages and new volunteers during experimental period who did not receive an email message (non-cohort). Percent with more than one session is the count of volunteers with more than one session divided by the number of contributors from Table 1. No differences are significant.

Message Pairs	**Difference**	χ^2	**Unadjusted p-value**	**Confidence Interval (99.15%)**
Contribution to Science—Science Learning	9.0%	3.0811	0.0792	-4.2%, 22.2%
Science Learning-Community	5.4%	0.9166	0.3384	-8.5%, 19.2%
Altruism—Science Learning	7.9%	2.3499	0.1253	-5.2%, 20.9%
Contribution to Science-Community	3.6%	0.3839	0.5355	-10.1%, 17.4%
Altruism—Contribution to Science	1.2%	0.0226	0.8806	-14.1%, 11.8%
Altruism—Community	2.5%	0.1544	0.6944	-11.1%, 16.1%

Table 5. Results of tests comparing the proportion of contributors who made contributions in more than one session between each pair of message conditions. No differences are significant.

As is expected, the distribution of the number of sessions per volunteer is quite skewed (most people have only one session but a few have a lot), as indicated by the difference between the mean and the median values and the high standard deviation. We therefore tested whether there was a difference between the cohorts in the number of sessions per volunteer with a non-parametric Kruskal–Wallis test. The test showed that there is no statistically significant difference among the per volunteer number of sessions across the cohorts (χ^2 (3, N=808) = 3.95, p = 0.2663).

We ran the same proportion test on the fraction of the volunteers who became sustained contributors (the Contributors percentage column) with the same Bonferroni correction. The results are shown in Table 5. The results show that the proportion of users who sustain their contribution is highest for *Contribute to Science* and *Altruism*, followed by *Community*, with *Learning* at the bottom. However, none of the differences are statistically significant. Note though that the confidence intervals are broad (about 25%), more than twice the greatest difference. The wide intervals suggest that the tests suffer from a lack of power to resolve the differences seen and that with a larger sample, the differences could be significant.

Cohort vs. non-cohort contributions. Finally, as noted above, we collected data on all new volunteers who joined the project after the message was sent. We considered the possibility that Zooniverse members who are on the mailing list might differ from those who are not on the list in their interest in contributing. To test this possibility, we compared the count of sessions from volunteers who received the recruiting messages to the count for those who did not receive the messages. The result of Wilcoxon signed-rank test indicated that contributions of volunteers who did not receive the recruiting message are significantly different from the ones who did (W=1228500, p = 0.00003).

We were concerned that the non-cohort sample might differ from the cohort sample because of the timing of when they joined. To address this concern, we first compared the distribution of the number of new volunteers vs. their start date in each cohort and the non-cohort. We found the shapes of the curves to be roughly similar, with a peak of new members at the project announcement, dropping off steadily afterwards. Though activity did not drop off completely in either group (e.g., there were volunteers who received an email in October who made their first contribution at the end of January), there were proportionally more non-cohort volunteers joining further after the announcement than from the cohorts. It could be that these late-joining non-cohort members simply have had less time to contribute, not less interest in contributing.

To check if this late activity was biasing the results, we computed a weighted average of the number of classifications and sessions per volunteer in the non-cohort, giving more weight to the earlier contributions and less to the later ones, so that the distribution of volunteers over time matched. To our surprise, this process actually made the differences between the cohort and non-cohort groups bigger. Apparently the earlier non-cohort contributors actually contributed less than later ones, despite having had more time in which to contribute. In retrospect, it is not surprising that the timing has little effect on the results. The majority of volunteers contribute for only one day, so the timing of data collection has little effect.

Question 2: Which message attracted the highest number of contributions from volunteers?

We answered the second question in two ways, looking first at the average number of contributions from volunteers in each cohort and then considering the contributions from the cohort as a group.

Average number of contributions

The Contributions columns of Table 6 gives the total number of classifications done by members of each cohort, the average and median number of classifications per volunteer and the standard deviation. As should be expected, the distribution of the number of contributions per volunteer are quite skewed—most people contribute only a few classifications and a few contribute a lot—as indicated by the difference between the mean and the median values and the high standard deviation. Figure 1 shows the distribution of classifications done per volunteer in the four cohorts using violin plots. A violin plot is like a box plot, but includes a kernel density plot for the data, thus showing the distribution in more detail. Note that the y-axis is log transformed to correct for the skew.

Cohort	**Classifications done**			
	Total	**Mean**	**Med.**	**SD**
Altruism	53,321	239.1	46	661.1
Contribution to Science	63,151	293.7	58	795.5
Community	30,844	175.3	47.5	415.9
Science Learning	38,140	196.6	43	523.9
Total of 4 cohorts	185,456	229.5	49	627.0
Non-cohort	520,972	185.5	40	658.4

Table 6. Contribution statistics for experimental groups: 4 cohorts who received messages and new volunteers during experimental period who did not receive an email message (non-cohort).

As the count of contributions per volunteer is not normally distributed, we tested whether there was a difference between the cohorts with a non-parametric Kruskal–Wallis test. The test showed that there is no statistically significant difference among the per volunteer count of contributions across the cohorts (χ^2 (3, N=808) = 1.378, p = 0.71). In summary, although volunteers in the *Contribute to Science* cohort did more classifications in comparison to the others, because of the high variability in contributions among volunteers within a cohort, none of the cohorts is statistically significantly different from the rest on the number of contributions per volunteer. The confidence intervals for the pair-wise tests have a range of about 25 to 30, which is quite a bit more than the differences. The wide intervals suggest that the test may suffer from a lack of power to resolve the differences seen and that with a larger sample, the differences could be significant. *Cohort vs. non-cohort contributions*. As above, we compared the count of classifications of volunteers who received the recruiting messages to the count for those who did not receive the messages. The result of Wilcoxon signed-rank test indicated that volunteers who did not receive a recruiting message made significantly fewer contributions to the project than did the volunteers who received and responded to the recruiting message (W= 11890, $p = 0.0367$).

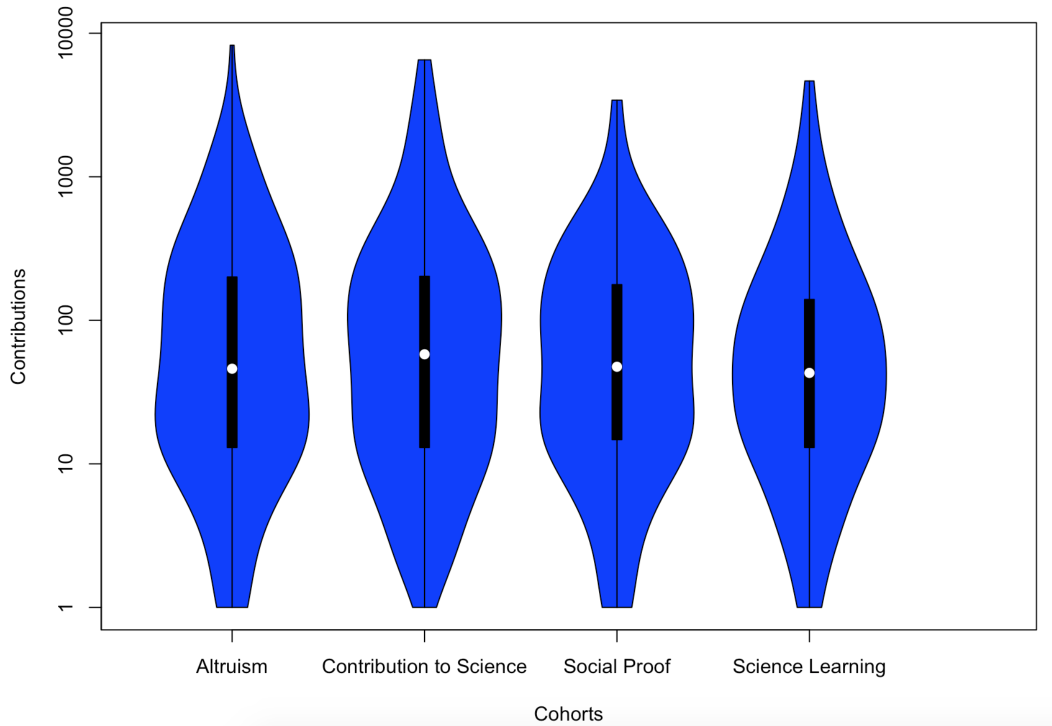

Figure 1. Violin plot of contributions per user by cohort, on log axis.

Total contributions

Finally, we examined the total number of contributions provided by each cohort, which is the combined result of attracting more volunteers and attracting volunteers who contribute more (or motivating volunteers to contribute more). Table 6 shows that *Contribute to science* led to the most total contributions being contributed, more than double the count for *Community* However, this difference could be due to chance. Recall that the volunteers in *Contribute to science* provided on average about 67% more classification each than those in *Community* but the high variability within cohorts meant that the difference was not statistically significant.

To test whether the total contributions received from a cohort is more or less than could be expected by chance requires knowing the distribution for total contributions. However, we do not have a sample of cohorts from which to determine this distribution empirically (as we did for average number of classification per user). To address this lacuna, we generated a set of random cohorts from the data for the actual respondents. We created a random cohort by randomly assigning each of the volunteers to one of four cohorts. This process randomly varied the cohorts along the two differences among cohorts we discussed above: how many volunteers are in the cohort and how many contributions the participants in the cohort make. Following this process, we created 1000 random cohorts of varied sizes and with varying samples of volunteers and so varying numbers of total contributions. To avoid creating correlations among the cohorts, each time we generated random cohorts we kept only one of the four. The plot of the distribution of the total number of contributions in the resulting sample of random cohorts is shown in Figure 2.

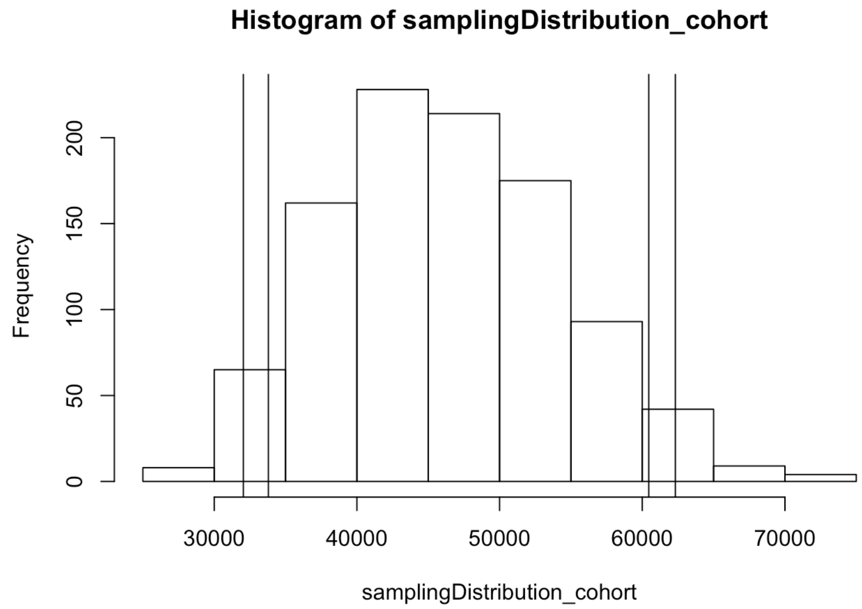

Figure 2. Distribution of total contributions for 1000 randomly generated cohorts drawing from mailing list respondents (all cohorts). Vertical lines indicate 5% and 10% upper and lower bounds.

Once we had a set of random cohorts, we could test whether the observed counts of total contributions are different that could be expected by chance by simply noting where in this distribution the actually-observed cohorts fall. This analysis shows that the total number of contributions received in *Contribute to science* (63,151) is at the 98th percentile (i.e., it is greater than 98% of the randomly generated cohorts), while the percentile for *Community* (30,844) is at the 2nd (i.e., smaller than 98% of the random cohorts). In other words, the total received in these two cohorts are respectively more and less than one would expected by a chance arrangement of the volunteers into cohorts, at $p<0.05$, suggesting that *Contribute to science* was particularly good at attracting contributions and *Community* was particularly poor.

Summary of findings

Table 7 provides a summary of the findings of this study. Our experiment shows that a message appealing to the motivation of *Contribution to science* worked better than three others in attracting volunteers to a citizen science project, and even though the average number of contributions per user is not statistically significantly greater, the total volume of contributions received in response to this message is greater than can be explained by chance. In contrast, *Community* message, while receiving a similar number of click throughs to other messages, had a lower level of overall contribution than expected by chance.

Discussion

While the experimental design provides good assurance about the results, it does not help to explicate the underlying mechanism for the results, that is, why does an appeal to science work better and an appeal to community worse? Specifically, it does not address the question of whether the results are due to selection, meaning that the message attracts participation by individuals with particular motivations who do more, or whether it makes salient a motivation that encourages more contribution by the recipients.

	Participation			Contribution	
Cohort	**Decide**	**Initial**	**Sustained**	**Average**	**Total**
Altruism	–	+			
Contribution to Science	+		(+)	(+)	+
Community				(–)	–
Science Learning	+		(–)		

Table 7. Summary of findings. + = significantly higher, – = significantly lower, (+/–) = higher or lower, though not statistically significant.

In this regard, the difference between the click through and participation percentage for *Community* is illuminating. Recall that participants receiving this message clicked through the site at about the same rate as others, but contributed at a much lower rate. A possible explanation for these findings is the nature of interaction in the projects. Recall that community engagement was a significant motivator in Foldit [Curtis, 2015]. In this system, volunteers do interact with other citizen scientists; thus, they could perceive that participating in the project as a sort of community. In contrast, there is not much visible community on Zooniverse sites. A new volunteer would need to explore the site to find group discussions and may need considerable expertise at the task to be able to contribute to the discussion. So, it could be that the message attracted volunteers interested in joining a community who were disappointed by the apparent lack of community when they first visited the site, leading to lower contribution.

The difference between selection and motivation as a mechanism connecting the messages and participation has important implications for how our results might influence practice. If the question is simply which message is more motivating, then a project manager can simply pick an appeal to science. If, however, different messages are encouraging different segments of the population to respond (selection), then potentially all messages could help by attracting different population segments. Note that the response rates to the individual messages are all quite low, so being able to add the responses together by attracting different segments would be a substantial improvement in the response. Further, if different volunteers are motivated by different factors, appeals should be tailored to the volunteer. Since the Zooniverse mails project announcements to volunteers regularly, a possible strategy is to try different strategies until one attracts a particular volunteer, and then to try that appeal again in future messages to that volunteer.

A second finding is that the efficacy of different motives does seem to change over time. Specifically, *Learning science* was as good at attracting click throughs and initial participation but seemed to fare worse in attracting sustained contribution. It could be that volunteers who were motivated by the opportunity to learn about a new branch of science had that interest fulfilled by their interaction with the project tutorials and science materials and so did not feel a need to continue to work on the glitch classification task, which is only tangentially related to the science of gravitational waves.

A further finding of the study is that the volunteers who responded to the recruiting message contributed significantly more than volunteers who joined about the same time but without having received a message. Again, the implication of this finding depends on whether the message is motivating or selecting volunteers. From the former perspective, the messages are doing what they should in encouraging participation. But from the later perspective, preferred above, it should not be surprising that volunteers who signed up for the mailing list are simply more motivated than those who did not, the content of the message notwithstanding. In either case, this finding emphasizes the importance of reaching out to prospective volunteers in multiple ways, and to consider channels for reaching and motivating different groups of volunteers.

Finally, the data are consistent with prior theorizing that notes that motivations for initial and sustained participation are different. While an appeal to altruism did worse in attracting visitors to the site, those visitors went on to contribute at a higher level. This finding suggests that messages at different points in a volunteer's engagement with the project might appeal to different motives: one set of motives to get a prospective volunteer to visit the site (e.g., appeal to learning), another to convince them to try it (e.g., appeal to altruism), and third to promote sustained contribution (e.g., contribution to science).

Study limitations

The design of the study reported in this paper is a true experiment, which addresses many threats to internal validity. However, there are some threats to construct validity. First, message recipients do not have to click on the link provided in the email message to access the system, so the click through rate might be an underestimate of the true interest. Conversely, a volunteer might forward the message to a friend who clicks the custom link. However, we also have counts of actual participation that are not affected by this problem. A second threat is to statistical conclusion validity. It appears that some of the statistical tests are underpowered, so some negative results could be different with a larger sample.

While experiments provide good internal validity, this validity comes at cost of possible threats to external validity. First, we only tested four specific versions of the messages. It could be that slight tweaks to the messages would change their performance, and we know nothing about the performance of appeals to other motives. It might even be possible to craft messages that combine aspect of different motivations, thus appealing to multiple segments of the population at once. Second, we ran the experiment in only one single project. Prior research on motivation has noted the importance of interest in the science, so projects with different science presumably attract different participants. It would be interesting to know if same results hold in other citizen science projects. Finally, given the particular nature of motivation for citizen science projects, we would not expect the finding to hold in other online communities, though some of the broader implications might (e.g., the evolution of motivations with participation).

Conclusion

The experiment reported here has both theoretical and practical implications. First, the work extends prior work on reported motivations by showing how these motivations work as an appeal to initial volunteers to a citizen science project. Specifically, our results provide further evidence for the importance of making a real contribution to science as a motivation for citizen science participants.

Practically, the work provides guidance to those who run citizen science projects. We examined three different outcomes and show that depending on the goal of recruitment, different messages may be more or less effective. In particular, if the goal is increasing the number of participants who are aware of the project, then appealing to the chance to contribute to or to learn about science seems to attract more visits than an appeal to altruism, though the later is more successful in contributing volunteers to contributors. And over all, an appeal to the chance to contribute to science seems to result in the largest number of contributions to the project. In summary, our results show that at least for the Zooniverse, citizen science projects are science, and that is reflected in the success of that message.

References

Bonney, R., Ballard, H., Jordan, R., McCallie, E., Phillips, T., Shirk, J., & Wilderman, C. C. (2009). Public participation in scientific research: Defining the field and assessing its potential for informal science education *A CAISE Inquiry Group Report*. Washington, DC: Center for Advancement of Informal Science Education (CAISE). Available from: https://eric.ed.gov/?id=ED519688.

Brossard, D., Lewenstein, B., & Bonney, R. (2005). 'Scientific knowledge and attitude change: The impact of a citizen science project'. *International Journal of Science Education, 27*(9), 1099–1121. doi: 10.1080/09500690500069483.

Cialdini, R. B. (2001). 'Harnessing the science of persuasion'. *Harvard Business Review, 79*(9), 72–81.

Cox, J., Oh, E. Y., Simmons, B., Graham, G., Greenhill, A., Lintott, C., Masters, K., & Woodcock, J. (2017). 'Doing good online: The changing relationships between motivations, activity and retention among online volunteers'. *Nonprofit and Voluntary Sector Quarterly*.

Crall, A., Kosmala, M., Cheng, R., Brier, J., Cavalier, D., Henderson, S., & Richardson, A. D. (2017). 'Volunteer recruitment and retention in online citizen science projects using marketing strategies: Lessons from Season Spotter'. *JCOM: Journal of Science Communication, 16*(1), A1.

Crowston, K., & Fagnot, I. (2008, 1–2 July). The motivational arc of massive virtual collaboration. In *Proceedings of the IFIP WG 9.5 Working Conference on Virtuality and Society: Massive Virtual Communities*, Lüneberg, Germany.

Curtis, V. (2015). 'Motivation to participate in an online citizen science game: A study of Foldit'. *Science Communication, 37*(6). doi: 10.1177/1075547015609322.

Domroese, M. C., & Johnson, E. A. (2017). 'Why watch bees? Motivations of citizen science volunteers in the Great Pollinator Project'. *Biological Conservation, 208*, 40-47. doi: 10.1016/j.biocon.2016.08.020.

Gugliucci, N., Gay, P., & Bracey, G. (2014). Citizen science motivations as discovered with CosmoQuest. In *Proceedings of the Ensuring Stem Literacy: A National Conference on STEM Education and Public Outreach*, pp. 437–440.

Holohan, A., & Garg, A. (2005). 'Collaboration online: The example of distributed computing'. *Journal of Computer-Mediated Communication, 10*(4). doi: 10.1111/j.1083-6101.2005.tb00279.x.

Kaufman, G., Flanagan, M., & Punjasthitkul, S. (2016). Investigating the impact of 'emphasis frames' and social loafing on player motivation and performance in a crowdsourcing game. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*, Santa Clara, CA, USA, pp. 4122-4128. doi: 10.1145/2858036.2858588.

Kraut, R. E., & Resnick, P. (Eds.). (2011). *Building successful online communities: Evidence-based social design*. Cambridge, MA: MIT Press.

Land-Zandstra, A. M., Devilee, J. L. A., Snik, F., Buurmeijer, F., & Broek, J. M. v. d. (2016a). 'Citizen science on a smartphone: Participants' motivations and learning'. *Public Understanding of Science, 25*(1), 45-60. doi: 10.1177/0963662515602406.

Land-Zandstra, A. M., van Beusekom, M., Koppeschaar, C., & van den Broek, J. (2016b). 'Motivation and learning impact of Dutch flu-trackers'. *Journal of Science Communication, 15*(1), A04. Available from: : https://jcom.sissa.it/archive/15/01/JCOM_1501_2016_A04

Mao, A., Kamar, E., & Horvitz, E. (2013). Why stop now? Predicting worker engagement in online crowdsourcing. In *Proceedings of the Conference on Human Computation and Crowdsourcing*, Palm Springs, CA, USA.

Masters, K., Oh, E. Y., Cox, J., Simmons, B., Lintott, C., Graham, G., Greenhill, A., & Holmes, K. (2016). 'Science learning via participation in online citizen science'. *Journal of Science Communication, 15*(3), A07. Available from: https://jcom.sissa.it/archive/15/03/JCOM_1503_2016_A07

McInnis, B. J., Murnane, E. L., Epstein, D., Cosley, D., & Leshed, G. (2016). One and done: Factors affecting one-time contributors to ad-hoc online communities. In *Proceedings of the ACM Conference on Computer-Supported Cooperative Work & Social Computing*, pp. 609-623.

Nov, O., Arazy, O., & Anderson, D. (2011). Dusting for science: Motivation and participation of digital citizen science volunteers. In *Proceedings of the iConference*, Seattle, Washington, USA, pp. 68-74. doi: 10.1145/1940761.1940771.

Oreg, S., & Nov, O. (2008). 'Exploring motivations for contributing to open source initiatives: The roles of contribution context and personal values'. *Computers in Human Behavior, 24*(5), 2055-2073. doi: 10.1016/j.chb.2007.09.007.

Raddick, M. J., Bracey, G., Gay, P. L., Lintott, C. J., Murray, P., Schawinski, K., Szalay, A. S., & Vandenberg, J. (2010). 'Galaxy Zoo: Exploring the motivations of citizen science volunteers'. *Astronomy Education Review, 9*(1), 010103–010118. doi: 10.3847/AER2009036.

Reed, J., Raddick, M. J., Lardner, A., & Carney, K. (2013, 7-10 Jan. 2013). An exploratory factor analysis of motivations for participating in Zooniverse, a collection of virtual citizen science projects. In *Proceedings of the Hawaii International Conference on System Sciences*, pp. 610-619. doi: 10.1109/HICSS.2013.85.

Robson, C., Hearst, M., Kau, C., & Pierce, J. (2013). Comparing the use of social networking and traditional media channels for promoting citizen science. In *Proceedings of the Conference on Computer-supported Cooperative Work*, San Antonio, Texas, USA, pp. 1463–1468. doi: 10.1145/2441776.2441941.

Rotman, D., Preece, J., Hammock, J., Procita, K., Hansen, D., Parr, C., Lewis, D., & Jacobs, D. (2012). Dynamic changes in motivation in collaborative citizen-science projects. In *Proceedings of the Conference on Computer Supported Cooperative Work*, Seattle, Washington, USA, pp. 217-226. doi: 10.1145/2145204.2145238.

Scheliga, K., Friesike, S., Puschmann, C., & Fecher, B. (2016). 'Setting up crowd science projects'. *Public Understanding of Science*, 1–20. doi: 10.1177/0963662516678514.

West, S., & Pateman, R. (2016). 'Recruiting and retaining participants in citizen science: What can be learned from the volunteering literature?'. *Citizen Science: Theory and Practice, 1*(2), 15. doi: 10.5334/cstp.8.

Wright, D. R., Underhill, L. G., Keene, M., & Knight, A. T. (2015). 'Understanding the motivations and satisfactions of volunteers to improve the effectiveness of citizen science programs'. *Society & Natural Resources, 28*(9), 1013–1029. doi: 10.1080/08941920.2015.1054976.

Zevin, M., Coughlin, S., Bahaadini, S., Besler, E., Rohani, N., Allen, S., Cabero, M., Crowston, K., Katsaggelos, A., Larson, S., Lee, T. K., Lintott, C., Littenberg, T., Lundgren, A., Oesterlund, C., Smith, J., Trouille, L., & Kalogera, V. (2017). 'Gravity Spy: Integrating Advanced LIGO detector characterization, machine learning, and citizen science'. *Classical and Quantum Gravity, 34*(6). doi: 10.1088/1361-6382/aa5cea.

Authors

Tae Kyoung Lee, PhD, is an Assistant Professor in the Department of Communication at the University of Utah. Her PhD is from Cornell University. Her research focuses on message processing and effects.

Kevin Crowston is a Distinguished Professor of Information Science in the School of Information Studies at Syracuse University. He received his Ph.D. (1991) in Information Technologies from the Sloan School of Management, Massachusetts Institute of Technology (MIT). His research examines new ways of organizing made possible by the extensive use of information and communications technology. Specific research topics include the development practices of Free/Libre Open Source Software teams and work practices and technology support for citizen science research projects.

Mahboobeh (Mabi) Harandi is a PhD student in the School of Information Studies (iSchool) at Syracuse University. She got her MSc in information systems from Norwegian University of Science and Technology where she was involved in the SmartMedia program and studied machine learning techniques and sentiment analysis over news articles. Her current research is studying user behavior/experience in online communities. She applies different approaches of natural language processing and statistical test to analyze the user behaviors.

Carsten Østerlund is an associate professor at the School of Information Studies, Syracuse University. He received his Ph.D. from MIT's Sloan School of Management (2003) and a M.A. in social psychology and social anthropology from University of Aarhus and University of Copenhagen, Denmark. During his MA studies he spent two years as a Fulbright scholar at UC Berkeley, Department of Social and Cultural Studies. His interests include distributed & virtual work, organizational learning and knowledge, communication practices, and medical informatics.

Grant Miller was awarded a PhD in astrophysics in 2013, working on the detection and characterisation of transiting exoplanets at the University of St Andrews. He then joined the Zooniverse, the world's leading citizen science research group, at the University of Oxford where he is now project manager and communications lead.

Appendix: Full text of recruiting email messages

Condition 1. Science learning

Subject: Gravity Spy: Extend your knowledge of astrophysics!

Hi there,

I'm thrilled to tell you about a brand new Zooniverse project - Gravity Spy

On September 14th 2015, a century after Einstein predicted the existence of ripples in spacetime known as gravitational waves, the Laser Interferometer Gravitational Wave Observatory (LIGO) made the first direct detection of this elusive phenomenon.

Being the most sensitive and most complicated gravitational experiment ever created, LIGO is susceptible to a variety of non-cosmic artifacts known as glitches. By selecting the right classification for a given glitch, you can teach computers to do this classification themselves on much larger datasets.

In this project, you can learn how to identify all of the glitch morphologies and open up an even bigger window into the gravitational wave universe.

Get involved now at www.gravityspy.org

Condition 2: Community

Subject: Gravity Spy: Join your fellow citizen scientists!

Hi there,

I'm thrilled to tell you about a brand new Zooniverse project - Gravity Spy

Join your fellow citizen scientists in classifying problematic noise in the search for gravitational waves!

On September 14th 2015, a century after Einstein predicted the existence of ripples in spacetime known as gravitational waves, the Laser Interferometer Gravitational Wave Observatory (LIGO) made the first direct detection of this elusive phenomenon.

Being the most sensitive and most complicated gravitational experiment ever created, LIGO is susceptible to a variety of non-cosmic artifacts known as glitches.By selecting the right classification for a given glitch, you can teach computers to do this classification themselves on much larger datasets.

Many citizen scientists are already participating in the project, identifying all of the glitch morphologies and opening up an even bigger window into the gravitational wave universe.

Get involved now at www.gravityspy.org

Condition 3: Contribution to science

Subject: Gravity Spy: Contribute to Science!

Hi there,

I'm thrilled to tell you about a brand new Zooniverse project - Gravity Spy

You can contribute to science by classifying problematic noise in the search for gravitational waves!

On September 14th 2015, a century after Einstein predicted the existence of ripples in spacetime known as gravitational waves, the Laser Interferometer Gravitational Wave Observatory (LIGO) made the first direct detection of this elusive phenomenon.

Being the most sensitive and most complicated gravitational experiment ever created, LIGO is susceptible to a variety of non-cosmic artifacts known as glitches.

By selecting the right classification for a given glitch, you can teach computers to do this classification themselves on much larger datasets.

Through the Gravity Spy project, you can contribute to science, identify all of the glitch morphologies, and open up an even bigger window into the gravitational wave universe.

Get involved now at www.gravityspy.org

Condition 4: Altruism

Subject: Gravity Spy: Please help scientists!

Hi there,

I'm thrilled to tell you about a brand new Zooniverse project - Gravity Spy

Astrophysicists need your help to classify problematic noise in the search for gravitational waves!

On September 14th 2015, a century after Einstein predicted the existence of ripples in spacetime known as gravitational waves, the Laser Interferometer Gravitational Wave Observatory (LIGO) made the first direct detection of this elusive phenomenon. Being the most sensitive and most complicated gravitational experiment ever created, LIGO is susceptible to a variety of non-cosmic artifacts known as glitches.

By selecting the right classification for a given glitch, you can teach computers to do this classification themselves on much larger datasets. Through the Gravity Spy project, you can help scientists identify all of the glitch morphologies and open up an even bigger window into the gravitational wave universe!

Get involved now at www.gravityspy.org