

Everyone counts? Design considerations in online citizen science

Helen Spiers, Alexandra Swanson, Lucy Fortson, Brooke D. Simmons, Laura Trouille, Samantha Blickhan and Chris Lintott

Abstract

Effective classification of large datasets is a ubiquitous challenge across multiple knowledge domains. One solution gaining in popularity is to perform distributed data analysis via online citizen science platforms, such as the Zooniverse. The resulting growth in project numbers is increasing the need to improve understanding of the volunteer experience; as the sustainability of citizen science is dependent on our ability to design for engagement and usability. Here, we examine volunteer interaction with 63 projects, representing the most comprehensive collection of online citizen science project data gathered to date. Together, this analysis demonstrates how subtle project design changes can influence many facets of volunteer interaction, including when and how much volunteers interact, and, importantly, who participates. Our findings highlight the tension between designing for social good and broad community engagement, versus optimizing for scientific and analytical efficiency.

Keywords

Citizen science; Public engagement with science and technology; Social inclusion

DOI

<https://doi.org/10.22323/2.18010204>

Submitted: 3rd April 2018

Accepted: 20th November 2018

Published: 17th January 2019

Context

During the last decade, an increasing number of research teams have deployed online citizen science projects to aid with data analysis [Brabham, 2008]. Typically, these projects invite volunteers to complete a classification task associated with a single element of data, such as an image, graph or video clip, with multiple volunteers examining each separate data point. The growth of this mode of distributed data analysis is being driven by the increased availability of datasets in many research disciplines, coupled with the concurrent broad establishment and use of web-connected computer and mobile technology. In addition to being motivated by the need to produce annotated datasets for research purposes, project owners are frequently passionate about the potential of citizen science to engage the public in authentic research. This raises a range of interesting challenges and

questions for those studying, designing and implementing citizen science projects, such as how to effectively satisfy the dual aims of citizen science of scientific efficiency and social inclusivity. Here, we use a large dataset of more than 60 online citizen science projects, representing the most comprehensive collection of online citizen science project data gathered to date, to study how project design affects volunteer behaviour, and ultimately the success of the project.

We focus on a well-established platform for online citizen science, the Zooniverse. The Zooniverse is the largest and most popular citizen science platform for data interpretation [Woodcock et al., 2017] (Figure 1 and Figure 2). It has provided a space for volunteers to contribute to more than 100 distinct research projects across multiple domains, including astronomy, ecology, medicine and the humanities. Although diverse in subject matter, Zooniverse projects are unified by a common theme of asking volunteers to perform simple tasks such as image classification and text transcription, and the aggregation of these projects onto one platform confers a unique opportunity to examine volunteer behaviour across projects.

Citizen science involves the collaboration of the general public with professional scientists to conduct research. The principal benefit of applying this method is that it enables research that would not otherwise be possible; although computer-based analysis can address many research questions, it is yet to surpass human ability in a number of areas, including recognition of the often complex patterns that occur within data [Cooper et al., 2010; Kawrykow et al., 2012]. Other potential benefits to online citizen science based research include a reduction in data processing time and cost [Sauermann and Franzoni, 2015], and the engagement of a more diverse crowd that may include typically underrepresented skills or demographic features [Jeppesen and Lakhani, 2010; Pimm et al., 2014]. Online citizen science projects are an effective form of public engagement, providing volunteers with an opportunity to learn more about science [Masters et al., 2016] and with an easily accessible means to make an authentic contribution to a research project, which can improve public engagement with, and advocacy of, scientific material [Forrester et al., 2017; Straub, 2016].

Despite the relative infancy of online citizen science, it has already contributed many noteworthy discoveries across diverse research domains, including the identification of new types of galaxies [Cardamone et al., 2009], models of large carnivore coexistence in Africa [Swanson et al., 2016], elucidation of protein structures relevant to HIV transmission [Cooper et al., 2010; Khatib et al., 2011], classification of cancer pathology [Candido dos Reis et al., 2015] and mouse retinal connectome maps [Kim et al., 2014]. The growing success of citizen science, in conjunction with significant reductions in the barriers to developing online citizen science projects, has led to an exponential growth in the number and diversity of projects. This is creating both an opportunity, and a need, to further study and understand the volunteer experience [Cox et al., 2015; Sauermann and Franzoni, 2015], as the sustainability of citizen science is dependent on our ability to design for volunteer engagement while optimising for scientific efficiency.

There has already been much work to examine project design in the Zooniverse, though typically by researchers studying a small number of projects. For example, Jackson et al., investigated the effect of novelty on user motivation, showing that for one project (*Higgs Hunters*) a message informing volunteers when they were the

first to see a particular data object increased participation [Jackson, Crowston et al., 2016]. Lee et al., found that messages which refer to learning, contribution and social engagement were more effective than direct appeals to altruism [Lee et al., 2017]. Segal et al., used active messaging about 'helpfulness' in the longest running Zooniverse project, *Galaxy Zoo*, to further increase engagement [Lintott et al., 2008; Segal et al., 2016]. Sprinks et al., recently found that participants prefer task designs with greater variety and autonomy, however fulfilling this preference did not improve performance [Sprinks et al., 2017]. Together, these studies reveal that user motivations, as discerned by studies of behaviour, are predictably complex. The effect of participation on volunteers may be assumed as similarly complicated, though work by Masters et al., has shown that volunteers self-report learning both about the scientific topics with which they are engaged, and about the process of science itself [Masters et al., 2016]. This work also shows that learning about scientific topics occurs, and is correlated with further engagement in the project.

None of these studies has taken advantage of the large number of projects now available on the Zooniverse platform; completing such a survey would allow us to identify aspects of project design which have significant effects. Because Zooniverse projects use a common codebase they share many similar features (e.g. the flow from the landing page to the classification interface, the presence of discussion forums etc.), therefore, differences will be due to fundamental design choices, such as the difficulty of the task set, the choice of dataset and the amount of data available. In this study, we employ the Zooniverse online citizen science platform as a 'web observatory ecosystem' to study volunteer behaviour across $n = 63$ projects. This work extends preliminary analyses previously presented as a poster at The Web Conference 2018 [Spiers et al., 2018] through quantifying and comparing additional project measures, utilizing the similar number of astronomy and ecology projects to assess the academic domain specificity of our observations, and closely examining unique findings associated with an individual Zooniverse project, *Supernova Hunters*. The analyses presented here provide quantitative evidence and insight relevant to researchers designing and developing online citizen science projects, and are informative to researchers studying the crowd-based production of knowledge.

Methods

To conduct the data analysis performed in this study, we first assembled the most comprehensive collection of Zooniverse classification data to date; including data from $n = 63$ Zooniverse projects (Figure 1, Table 6). Although these projects are all hosted on the same platform, these projects differ in domain (Figure 2), launch date (Figure 1) and task, amongst other variables such as level of social media engagement and level of press attention. An overview of project characteristics is provided in Table 6.

Data for the $n = 63$ projects were obtained from Zooniverse databases. Projects were selected for inclusion in this analysis based on data availability. Project characteristics, including domain and launch date, are summarized in Table 6. Projects that have been rebuilt and relaunched were treated as separate projects for the purpose of this analysis (see *Notes from Nature*, *Milky Way Project* and *Chicago Wildlife Watch* in Table 6). Zooniverse projects can be highly variable in duration, from weeks to years, therefore, for ease of comparison across projects a consistent

Cumulative Zooniverse project count.

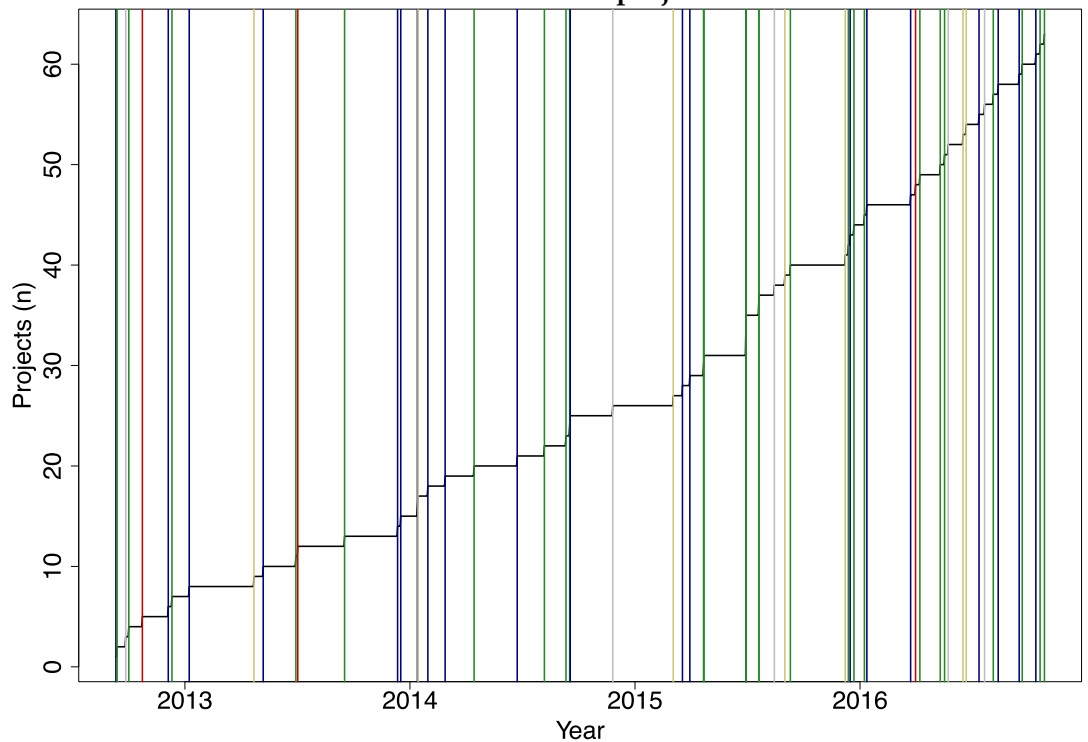


Figure 1. Accumulation of projects on the Zooniverse. The launch of the $n = 63$ projects included in this study are indicated by vertical lines. The line colour indicates the project domain (Ecology = green, Astronomy = blue, Transcription = yellow, Other = Grey, Biomedical = red). The black line shows the cumulative number of projects launched on the Zooniverse platform.

observation window of 100 days post-launch day was applied across all analyses. For logged-in, registered volunteers it is possible to describe volunteer-specific classification activity.

Throughout this paper, the term ‘classification’ is used to denote a single unit of analysis on a project by a volunteer, such as the tagging of an image or a video, whereas the term ‘subject’ refers to a single data object such as an image, video or graph. Different classification types can vary significantly in the amount of effort they demand. For a detailed glossary of Zooniverse terms, see Simpson et al., [Simpson, Page and De Roure, 2014]. For further information about the Zooniverse platform please see <https://www.zooniverse.org>.

To examine volunteer demographic features, data were extracted from Google Analytics (GA) (<https://analytics.google.com/>) for five astronomy and five ecology projects. To improve comparability between the projects examined and create a more uniform sample, we analysed the five most recently launched projects built using the project builder (<https://www.zooniverse.org/lab>) for both the astronomy (*Gravity Spy*, *Milky Way* (2016), *Radio Meteor Zoo*, *Supernova Hunters* and *Poppin’ Galaxy*) and ecology (*Chicago Wildlife Watch* (2016), *Arizona Batwatch*, *Mapping Change*, *Camera CATalogue*, *Snapshot Wisconsin*) domains, from our cohort of $n = 63$ projects. For each project examined, several variables were extracted from GA for the classify page of each project for the first quarter of 2017 (January 1st 2017

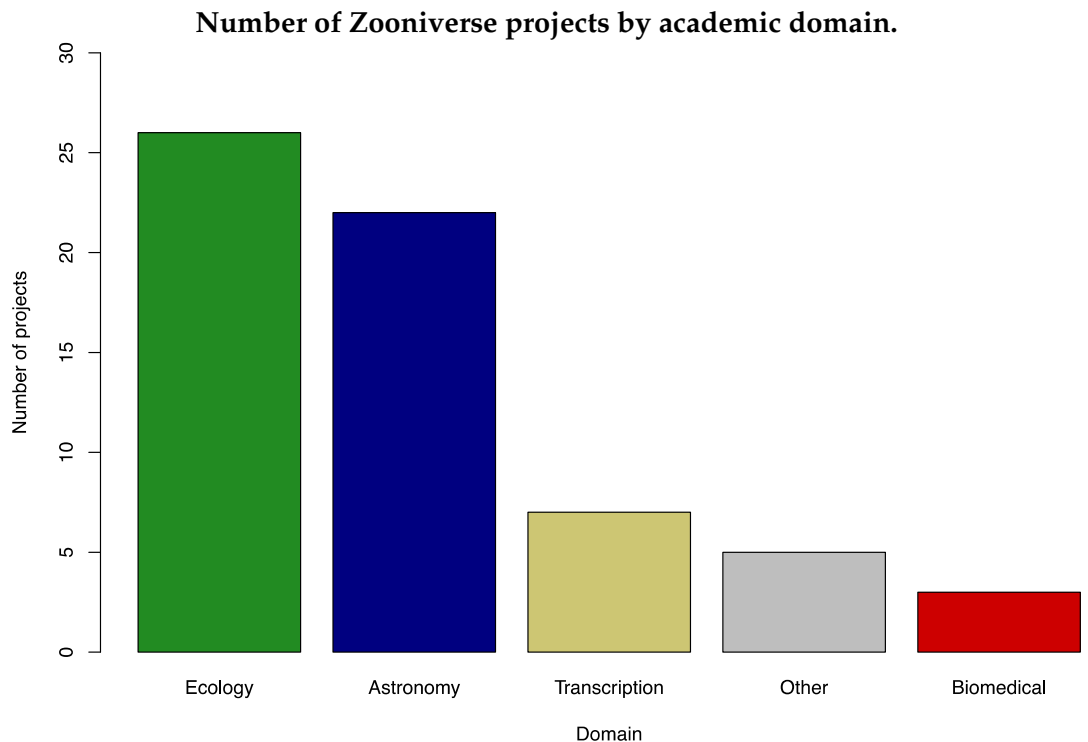


Figure 2. The Zooniverse platform supports a similar number of ecology and astronomy projects. Of the $n = 63$ projects included in this study; Ecology $n = 26$, Astronomy $n = 22$, Transcription $n = 7$, Other = 5, Biomedical = 3.

to March 31st 2017; data obtained May 9th 2017). These included the number of 'Page views' ('Page views is the total number of pages viewed. Repeated views of a single page are counted') subset by the secondary variables of age and sex.

The breadth and depth of the data collected by GA, in addition to it being a free and easily accessible service, has led to GA becoming an accepted research tool and one of the most frequently used methods to measure website performance [Clark, Nicholas and Jamali, 2014]. GA has been successfully used to examine user behaviour [Crutzen, Roosjen and Poelman, 2013; Song et al., 2018] website effectiveness [Plaza, 2011; Turner, 2010] and web traffic [Plaza, 2009]. However, it should be noted that it does have a number of constraints. For example, although GA uses multiple sources (third-party DoubleClick cookies, Android advertising ID and iOS identifier for advertisers [Google, 2018]) to extract user demographic information (age, gender and interests), these data may not be available for all users. In the analyses presented here, demographic data was available for 49.83% of total users for the variable of age, and for 53.41% of total users for the variable of gender. A further limitation is that occasionally the demographic data reported by GA may reflect a sample of website visitors, and hence may not be representative of overall site composition. Although the data presented in this study represents a sample of website visitors, for each page examined this sample was large (representing hundreds to thousands of page views) therefore the impact of missing demographic information from individual users or mis-sampled data will be minimal. The possibility of 'referrer spam' and 'fake traffic' pose further limitations to the accuracy of reports from GA, however these issues are less likely to influence non-commercial sites such as the Zooniverse.

Results

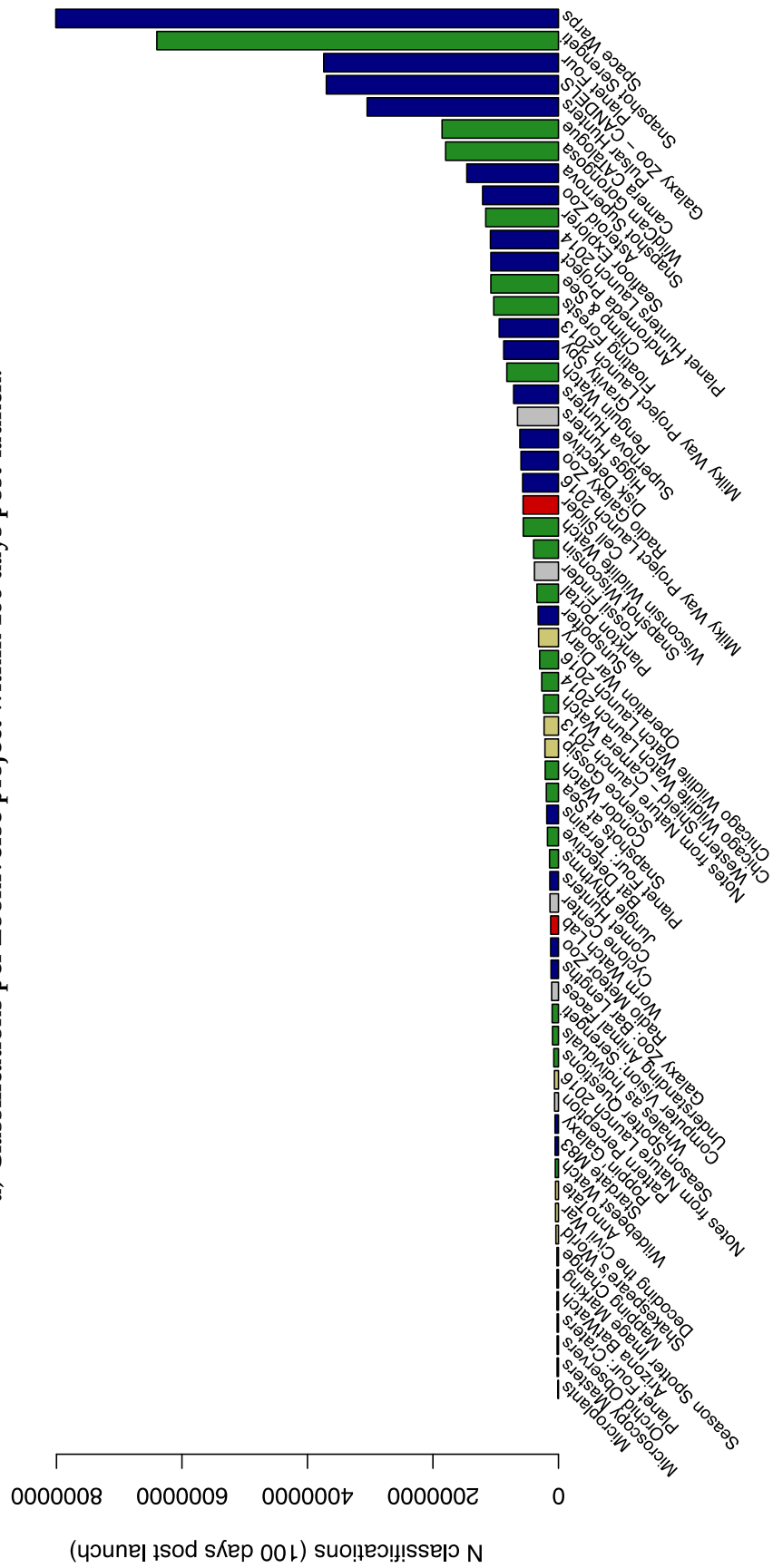
How heterogeneous is classification and volunteer activity across projects and academic domain?

A large amount of heterogeneity is found between the $n = 63$ Zooniverse projects for the total number of classifications received within the first 100 days post-launch (Figure 3a, Table 6). Notably, three orders of magnitude difference is observed between the project with the most classifications (*Space Warps*; total classifications $n = 8,011,730$) compared to the project with the fewest (*Microplants*; total classifications $n = 8,577$) (Table 6).¹ This suggests that the inclusion of a citizen science project on a successful citizen science platform website such as the Zooniverse does not guarantee high levels of engagement alone, as measured by number of classifications, and that some projects are far more successful at attracting classifications than others. Although initial inspection revealed a large difference between the number of classifications (within the first 100 days post-launch) for the $n = 26$ ecology projects (median = 224,054; interquartile range (IQR) = 78,928–758,447) compared to the $n = 22$ astronomy projects (median = 666,099; IQR = 150,642–1,178,327) (Table 1), this difference was not significant (P -value = 0.07, Mann-Whitney-Wilcoxon Test), indicating that other variables likely play an important role in the value of project metrics such as classification count.

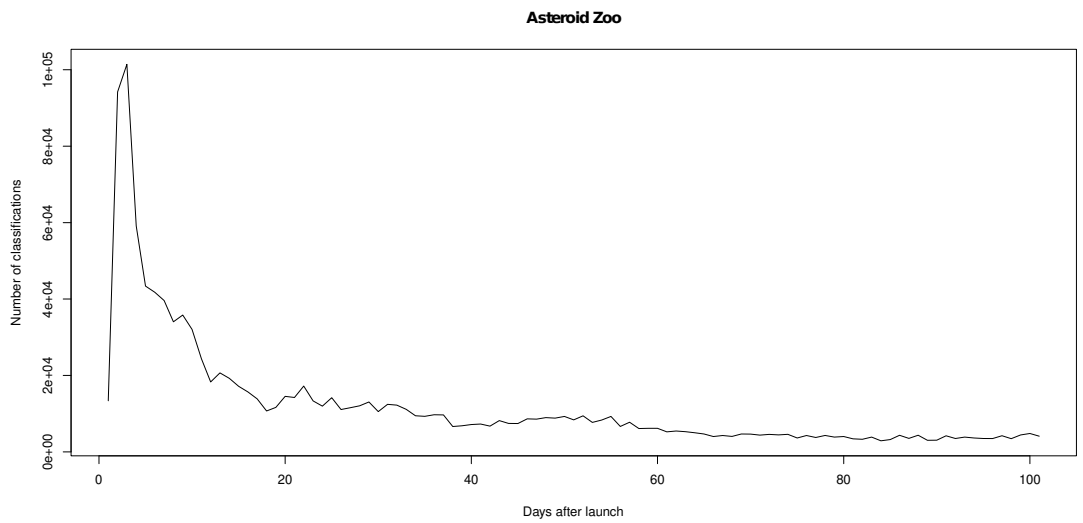
The absolute number of classifications received by each project may be influenced by factors such as how popular the project is or how much publicity it has received. Therefore, to examine whether volunteers experienced projects differently once contributing, we next analysed the median number of classifications made by registered volunteers within 100 days post-project launch for each of the 63 projects (Figure 3b, Table 6). Again, a large amount of heterogeneity was observed between projects, with a broad spectrum from the project with the highest median number of classifications per registered volunteer (*Pulsar Hunters*; $n = 100$), to the projects with the lowest (*Orchid Observers*, *Mapping Change* and *Decoding the Civil War*; $n = 3$), and this difference was highly statistically significant in each case (P -value = $< 2.2e-16$, Mann-Whitney-Wilcoxon Test). This indicates that once volunteers are participating in a project, a variety of engagement levels are observed. However, no significant difference was found (P -value = 0.47, Mann-Whitney-Wilcoxon Test) when comparing the median number of classifications per registered volunteer within the first 100 days post-launch for the $n = 26$ ecology projects (median = 15; IQR = 9–32) to the $n = 22$ astronomy projects (median = 17, IQR = 11–26).

¹The few classifications received by the *Microplants* project are likely due to this project being promoted primarily in a museum setting, as opposed to the broader Zooniverse community.

a) Classifications per Zooniverse project within 100 days post-launch.



a) Number of classifications a day on Asteroid Zoo.



b) Number of classifications a day on Supernova Hunters.

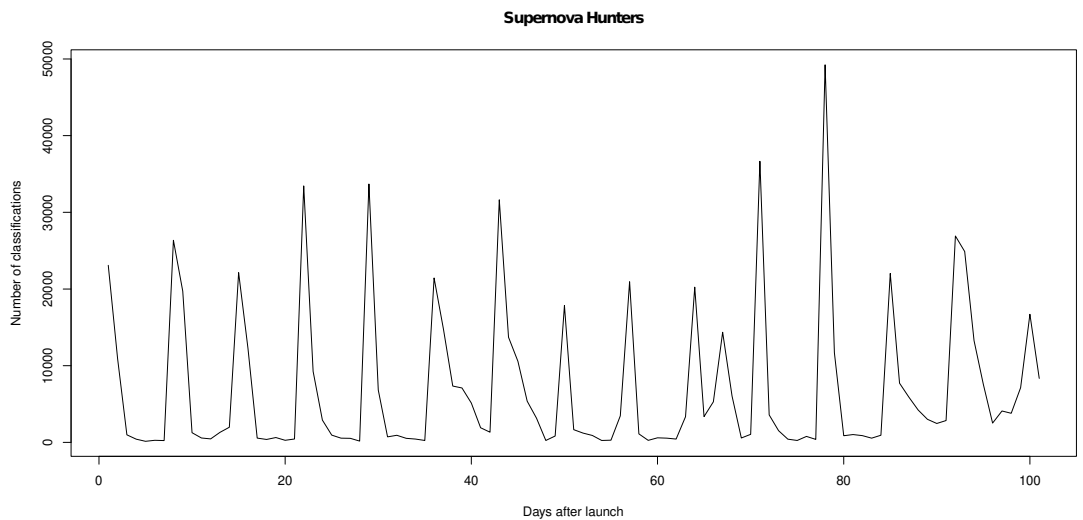


Figure 4. *Supernova Hunters* has a distinctive classification curve. A typical Zooniverse project has a classification curve displaying a peak of activity after launch that rapidly declines a), however there are exceptions to this observation, the most striking of which is the classification curve of the *Supernova Hunters* project b). Volunteers return regularly to this project upon a weekly release of new project data. This figure has been adapted from a poster presented at the World Wide Web Conference 2018 [Spiers et al., 2018]. Redistributed under the [Creative Commons Attribution 4.0 International \(CC BY 4.0\)](https://creativecommons.org/licenses/by/4.0/) license.

The majority of Zooniverse projects show similar temporal trends in classification activity — most projects possess a classification curve characterized by a high level of activity upon project launch (due to the sending of an e-newsletter to the Zooniverse volunteer community) that rapidly declines (Figure 4a), with intermittent spikes of activity, which can be the result of further project promotion, press coverage or the release of new data, amongst other factors. Of the 63 projects assessed, the *Supernova Hunters* project showed striking exception to this trend and instead had a classification curve displaying a recurring peak of activity each week (Figure 4b). This pattern of classification activity has arisen from *Supernova Hunters'* regular release of new project data and concurrent e-newsletter notification sent to project volunteers [Wright et al., 2017]. Notably, volunteer activity on the *Supernova*

Table 1. Classifications during the first 100 days post-launch, by domain. Median, first and third quartile of number of classifications received by project subset by academic domain, for the first 100 days post-launch.

| Domain | N projects | Median | First quartile | Third quartile |
|---------------|------------|--------|----------------|----------------|
| Astronomy | 22 | 666099 | 150642 | 1178327 |
| Ecology | 26 | 224054 | 78928 | 758447 |
| Biomedical | 3 | 123236 | 73320 | 344282 |
| Transcription | 7 | 63640 | 46092 | 222998 |
| Other | 5 | 135832 | 110051 | 383920 |

Hunters project has begun to precede the sending of e-newsletter notifications, suggesting that volunteers anticipate the release of new project data, and are therefore deeply engaged.

Table 2. Unique volunteers during the first 100 days post-launch, by domain. Median, first and third quartile of number of unique registered volunteers by project subset by academic domain, for the first 100 days post-launch.

| Domain | N projects | Median | First quartile | Third quartile |
|---------------|------------|--------|----------------|----------------|
| Astronomy | 22 | 3863 | 1760 | 5906 |
| Ecology | 26 | 1975 | 1289 | 3273 |
| Biomedical | 3 | 4177 | 2760 | 4644 |
| Transcription | 7 | 2300 | 1214 | 3838 |
| Other | 5 | 2730 | 1999 | 3016 |

The number of unique volunteers contributing to each project within the first 100 days post-launch, and its link to domain, was examined (Figure 3c). As with the number of classifications, the number of unique volunteers contributing to each project is highly heterogeneous (Table 2) and does not show any clear domain specificity; no significant difference (P -value = 0.07, Mann-Whitney-Wilcoxon Test) is observed between astronomy (median = 3863; IQR = 1760–5906) and ecology (median = 1975; IQR = 1289–3273) projects. As expected, the number of volunteers contributing to a project within the first 100 days post-launch is positively correlated with the number of classifications received by that project ($R = 0.64$, P -value = 2.30×10^{-8}). However, this correlation is not perfect, suggesting varying levels of volunteer contribution dependent on project. Rather than being stochastic, these differing volunteer contributions to individual projects are likely due to project specific variables. The heterogeneous contributions of volunteers to specific projects is also illustrated by the variable median number of classifications per project per day (Figure 5).

Median classifications a day for the first 100 days following project launch per project.

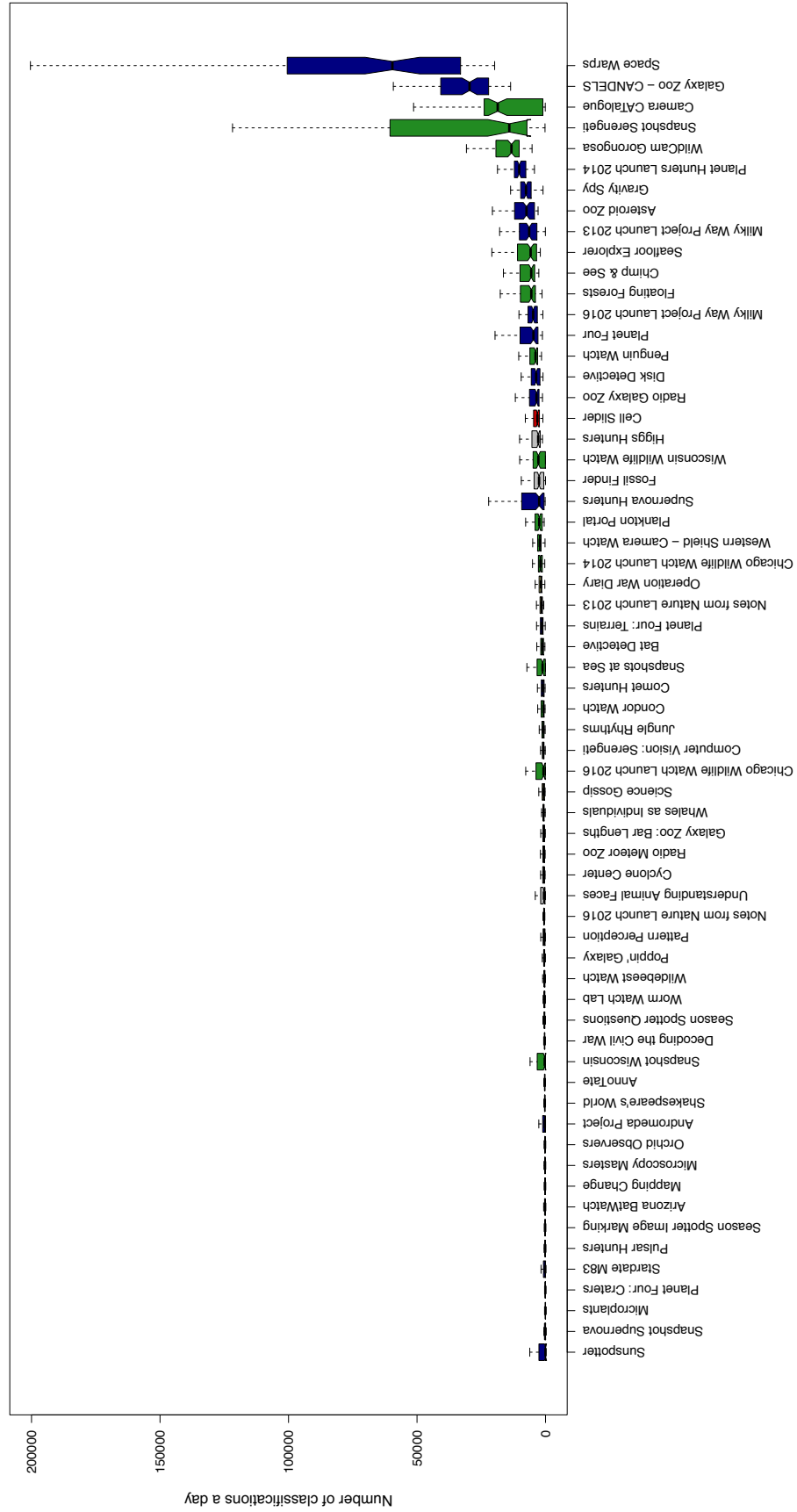


Figure 5. Zooniverse projects show heterogeneity in the number of classifications received a day. Projects are ordered by median classifications a day for the first 100 days following launch.

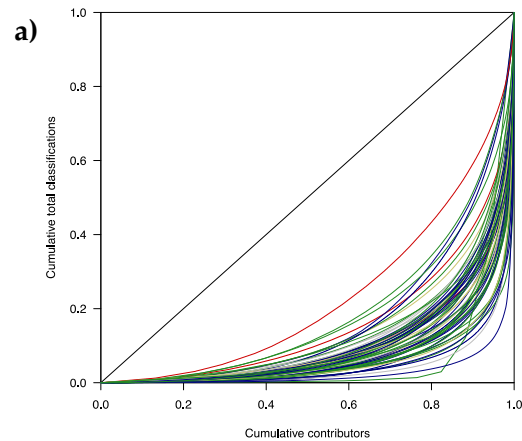
How skewed are volunteer contributions in Zooniverse projects?

Prior work has identified skewed volunteer contribution distributions in similar settings such as Wikipedia and OSS development [Franke and von Hippel, 2003; Ortega, Gonzalez-Barahona and Robles, 2008; Panciera, Halfaker and Terveen, 2009; Wilkinson, 2008] as well as Zooniverse projects [Cox et al., 2015; Sauermann and Franzoni, 2015]. We sought to extend previous analyses of Zooniverse projects by examining volunteer classification contribution inequality across a larger number of projects, which also enables assessment of our findings for domain specific trends. Applying an approach frequently used to examine income inequality [Gastwirth, 1972], we plotted the Lorenz curve for the distribution of volunteers' total classifications for each project (Figure 6a), and calculated corresponding Gini coefficients (Table 3).

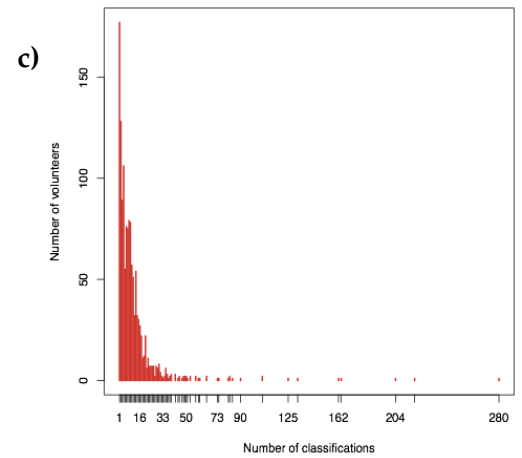
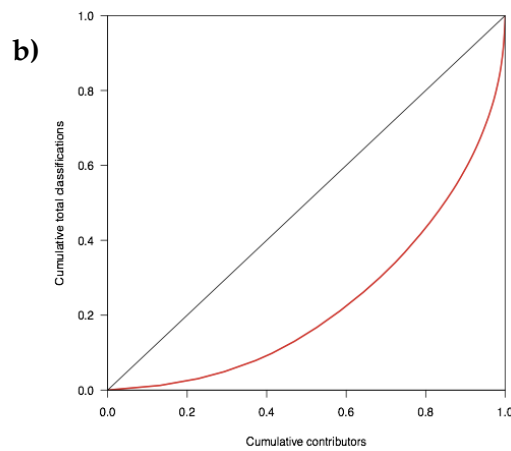
Table 3. *Project Gini coefficients.* Volunteer classification contribution inequality was assessed by calculating Gini coefficients for each project.

| Project_name | domain | Gini | Project_name | domain | Gini |
|------------------------------------|---------------|------|------------------------------------|---------------|------|
| Microscopy Masters | Biomedical | 0.54 | Shakespeare's World | Transcription | 0.8 |
| Season Spotter Image Marking | Ecology | 0.62 | Wisconsin Wildlife Watch | Ecology | 0.81 |
| Season Spotter Questions | Ecology | 0.65 | Andromeda Project | Astronomy | 0.81 |
| Planet Four: Craters | Astronomy | 0.65 | Asteroid Zoo | Astronomy | 0.82 |
| Pulsar Hunters | Astronomy | 0.68 | Sunspotter | Astronomy | 0.82 |
| Worm Watch Lab | Biomedical | 0.68 | Notes from Nature Launch 2013 | Transcription | 0.83 |
| Wildebeest Watch | Ecology | 0.7 | Higgs Hunters | Other | 0.83 |
| Science Gossip | Transcription | 0.72 | Camera CATalogue | Ecology | 0.83 |
| Understanding Animal Faces | Other | 0.75 | Mapping Change | Ecology | 0.83 |
| Computer Vision: Serengeti | Ecology | 0.75 | Snapshots at Sea | Ecology | 0.84 |
| Pattern Perception | Other | 0.75 | Operation War Diary | Transcription | 0.84 |
| Arizona BatWatch | Ecology | 0.75 | Galaxy Zoo — CANDELS | Astronomy | 0.84 |
| Chicago Wildlife Watch Launch 2016 | Ecology | 0.75 | Western Shield — Camera Watch | Ecology | 0.84 |
| Cyclone Center | Other | 0.76 | Chicago Wildlife Watch Launch 2014 | Ecology | 0.84 |
| Poppin' Galaxy | Astronomy | 0.76 | Comet Hunters | Astronomy | 0.84 |
| Whales as Individuals | Ecology | 0.77 | Planet Hunters Launch 2014 | Astronomy | 0.86 |
| Snapshot Supernova | Astronomy | 0.77 | Floating Forests | Ecology | 0.86 |
| AnnoTate | Transcription | 0.77 | Radio Galaxy Zoo | Astronomy | 0.86 |
| Snapshot Wisconsin | Ecology | 0.78 | WildCam Gorongosa | Ecology | 0.86 |
| Radio Meteor Zoo | Astronomy | 0.78 | Orchid Observers | Ecology | 0.86 |
| Snapshot Serengeti | Ecology | 0.78 | Milky Way Project Launch 2013 | Astronomy | 0.87 |
| Condor Watch | Ecology | 0.78 | Microplants | Ecology | 0.87 |
| Planet Four: Terrains | Astronomy | 0.78 | Notes from Nature Launch 2016 | Transcription | 0.88 |
| Seafloor Explorer | Ecology | 0.78 | Jungle Rhythms | Ecology | 0.88 |
| Planet Four | Astronomy | 0.79 | Disk Detective | Astronomy | 0.89 |
| Cell Slider | Biomedical | 0.79 | Plankton Portal | Ecology | 0.89 |
| Gravity Spy | Astronomy | 0.79 | Bat Detective | Ecology | 0.89 |
| Penguin Watch | Ecology | 0.79 | Milky Way Project Launch 2016 | Astronomy | 0.9 |
| Decoding the Civil War | Transcription | 0.8 | Space Warps | Astronomy | 0.9 |
| Galaxy Zoo: Bar Lengths | Astronomy | 0.8 | Fossil Finder | Other | 0.91 |
| Stardate M83 | Astronomy | 0.8 | Supernova Hunters | Astronomy | 0.94 |
| Chimp & See | Ecology | 0.8 | | | |

Lorenz curves for all Zooniverse projects.



Volunteer classification contributions in Microscopy Masters.



Volunteer classification contributions in Supernova Hunters.

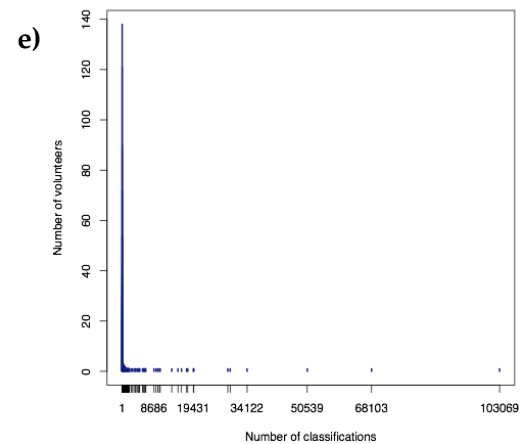
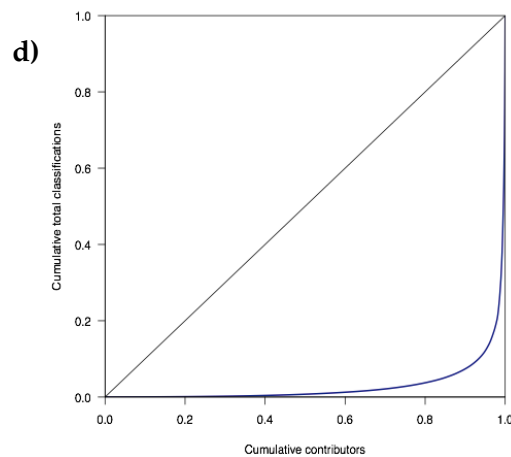


Figure 6. All Zooniverse projects display unequal volunteer classification contribution. a) Lorenz curves were plotted for all $n = 63$ projects to describe the inequality in number of classifications per registered volunteer, for the first 100 days post-launch. The plot shows the cumulative number of classifications versus the cumulative number of volunteers, with the increased curvature of the Lorenz curve indicating stronger inequality in volunteer contribution. The black 45° line corresponds to total equality, which in this case would represent all users contributing equal numbers of classifications. Although all projects displayed volunteer classification contribution inequality; a large amount of variation was observed between the project displaying the lowest degree of equality; *Microscopy Masters* (b, c), and the project displaying the highest; *Supernova Hunters* (d, e). Each plot is coloured by domain; Ecology = green, Astronomy = blue, Transcription = yellow, Other = Grey, Biomedical = Red. This figure has been adapted from a poster presented at the World Wide Web Conference 2018 [Spiers et al., 2018]. Redistributed under the [Creative Commons Attribution 4.0 International \(CC BY 4.0\)](https://creativecommons.org/licenses/by/4.0/) license.

Across all $n = 63$ projects a large area is observed between the Lorenz curve and the 45° line (perfect equality). This indicates that a large fraction of classifications are provided by a relatively small number of volunteers across all projects. However, heterogeneous volunteer contribution equality was identified between Zooniverse projects — the Gini coefficients varied from 0.54 for *Microscopy Masters* (Figure 6b, c), to 0.94 for *Supernova Hunters* (Figure 6d, e). Examining the mean Gini coefficient for each domain revealed similar Gini coefficients for ecology (0.80), astronomy (0.82), other (0.80) and transcription (0.81) projects, suggesting similar patterns of volunteer contribution across these domains.

Biomedical projects displayed a notably different average Gini coefficient of 0.67. This was significantly less than the Gini coefficient observed for astronomy projects (P -value = 0.046, Mann-Whitney-Wilcoxon Test), indicating a more unequal distribution of volunteer effort in astronomy projects compared to biomedical projects, suggesting that astronomy projects may attract a contingent of more prolific classifiers than biomedical projects. Notably, a comparably low Gini coefficient (0.72) has been reported elsewhere for the online biomedical citizen science project *Mark2Cure* [Tsueng et al., 2016]. This raises an interesting line of enquiry for future analyses — is there something about biomedical projects that disincentivises return volunteers, hindering the development of prolific classifiers, or are biomedical projects more successful in attracting a large number of casual contributors? Could this be related to perceived difficulty, or importance, of biomedical tasks? Further understanding patterns in volunteer contribution equality to different projects may provide insight regarding how to better design for the many or for the few, dependent on the particular project task.

Are Astronomy and Ecology projects associated with different volunteer demographics?

We next sought to describe basic demographic features of volunteers contributing to projects from differing domains. To perform these analyses, data were extracted from GA (<https://analytics.google.com/>) for the classification pages of five astronomy and five ecology projects (for a full description of approach, see the **Methods** section).

First, we examined the number of times the classification page was viewed for each project, subset by age of visitor for both astronomy (Figure 7a) and ecology projects (Figure 7b). No consistent trend was observed across the five astronomy projects examined (Figure 7a): for example, *Poppin' Galaxy* was more popular amongst the youngest age group (18–24) whereas *Supernova Hunters* is more popular amongst the oldest age group (65+), and other projects show no clear age bias in classification page views.

We see more uniformity in percentage page views by age group for the five ecology projects examined (Figure 7b). Although there is no striking overall trend, the peaks appear bimodal for the majority of projects examined, with peaks in percentage page views for 25–34 year olds and 55–64 year olds. The consistency in these peaks between the projects examined, particularly for the 55–64 year olds, indicates that ecology projects are more popular within these age groups. Such observations could be utilized to inform project promotion strategies.

Comparing the average page views count per age group for the five ecology projects to the five astronomy projects revealed a greater proportion of page views for astronomy projects relative to ecology projects for 18–24 year olds (Fisher’s Exact Test, odds ratio = 1.89, *P*-value = 8.86E-62) and 65+ year olds (Fisher’s Exact Test, odds ratio = 2.89, *P*-value = 4.16E-134). A smaller proportion of page views in astronomy projects relative to ecology projects was observed for the 25–34 year olds (Fisher’s Exact Test, odds ratio = 0.77, *P*-value = 5.67E-13), 45–54 year olds (Fisher’s Exact Test, odds ratio = 0.88, *P*-value = 5.11E-03), and a striking under enrichment in the 55–64 year olds (Fisher’s Exact Test, odds ratio = 0.39, *P*-value = 4.77E-156) (Table 4). These findings reflect the observations in Figure 7a and Figure 7b.

Table 4. *Page views by age group for ecology compared to astronomy projects.* Average page views for ecology compared to astronomy projects, subset by the demographic feature of age. Data were extracted from GA (see **Methods** section).

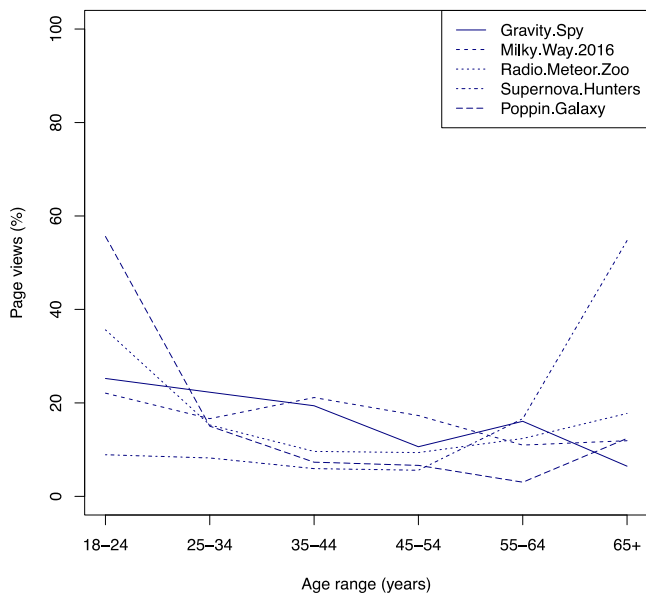
| Age range (years) | Ecology page views | % | Astronomy page views | % | Odds ratio | Lower 95% CI | Upper 95% CI | p value |
|-------------------|--------------------|-------|----------------------|-------|------------|--------------|--------------|-----------|
| 18–24 | 1242 | 13.11 | 2215 | 22.15 | 1.89 | 1.75 | 2.04 | 8.86E-62 |
| 25–34 | 1971 | 20.82 | 1677 | 16.78 | 0.77 | 0.71 | 0.82 | 5.67E-13 |
| 35–44 | 1504 | 15.89 | 1566 | 15.66 | 0.98 | 0.91 | 1.06 | 6.80E-01 |
| 45–54 | 1160 | 12.25 | 1096 | 10.96 | 0.88 | 0.81 | 0.96 | 5.11E-03 |
| 55–64 | 2826 | 29.85 | 1418 | 14.18 | 0.39 | 0.36 | 0.42 | 4.77E-156 |
| 65+ | 765 | 8.08 | 2025 | 20.26 | 2.89 | 2.64 | 3.16 | 4.16E-134 |

Next, we compared classification page views subset by sex for the five astronomy and five ecology projects, with data extracted from GA (see **Methods** section). A clear male bias can be observed in percentage classification page views across the five astronomy projects examined (Figure 7c), whereas ecology projects see more equality in percentage classification page views between the sexes (Figure 7d). Comparing the average page view counts by sex for the five ecology projects to the five astronomy projects revealed a highly significant greater proportion of males in astronomy projects compared to ecology projects (Fisher’s Exact Test, odds ratio = 3.92, *P*-value = 0), and a smaller proportion of females (Fisher’s Exact Test, odds ratio = 0.25, *P*-value = 0) (Table 5).

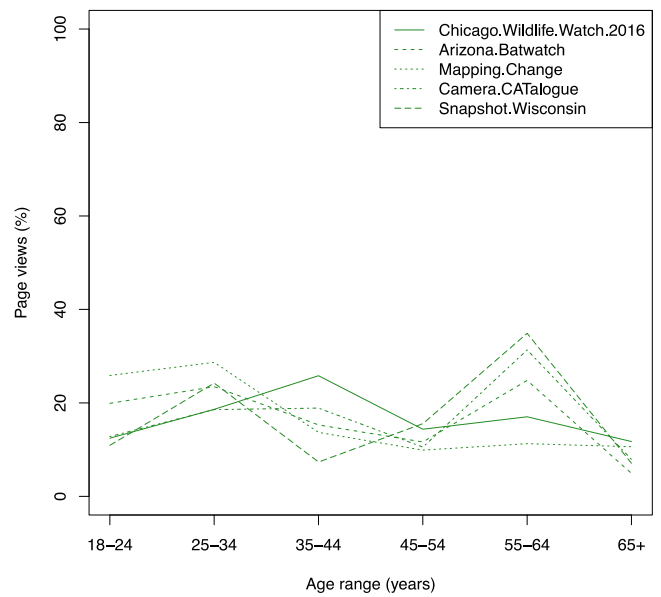
Table 5. *Page views by sex for ecology compared to astronomy projects.* Average page views for ecology compared to astronomy projects, subset by the demographic feature of Sex. Data were extracted from GA (see **Methods** section).

| Sex | Ecology project views | % | Astronomy project views | % | Odds ratio | Lower 95% CI | Upper 95% CI | p value |
|--------|-----------------------|-------|-------------------------|-------|------------|--------------|--------------|---------|
| Male | 4469 | 43.93 | 8127 | 75.45 | 3.92 | 3.7 | 4.16 | 0 |
| Female | 5703 | 56.07 | 2644 | 24.55 | 0.25 | 0.24 | 0.27 | 0 |

a) Page views by age for five astronomy projects.



b) Page views by age for five ecology projects.



c) Page views by sex for five astronomy and five ecology projects.

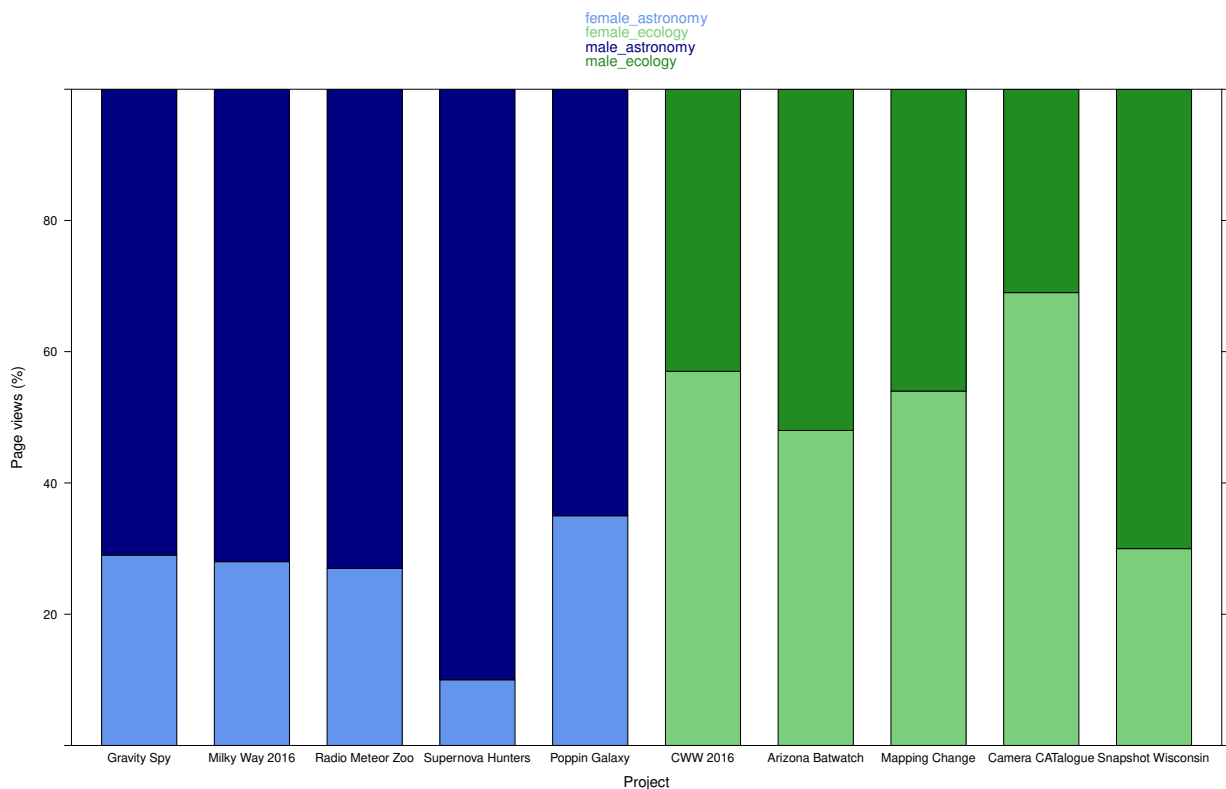


Figure 7. Domain-specific demographic features are observed for Zooniverse projects. Examining classification page views for astronomy and ecology projects, subset by age (a, b) and sex (c) revealed greater uniformity in page views by age group for ecology projects (b) compared to astronomy projects (a), and a clear male bias across astronomy projects (c). This figure has been adapted from a poster presented at the World Wide Web Conference 2018 [Spiers et al., 2018]. Redistributed under the [Creative Commons Attribution 4.0 International \(CC BY 4.0\)](https://creativecommons.org/licenses/by/4.0/) license.

Discussion

To ensure the sustainability of distributed data analysis via online citizen science, we must further understand the volunteer experience. In the analysis presented here we utilize a well-established online citizen science platform, the Zooniverse, as a 'web observatory ecosystem' to examine multiple project measures from the most comprehensive collection of online citizen science project data gathered to date.

Key findings include observation of a high level of heterogeneity, and little domain specificity, in the absolute number of classifications, number of unique volunteers and the median number of classifications per volunteer per project in the 100 days following project launch. These observations indicate that variables beyond the host citizen science platform and academic domain likely play an important role in project success. Such variables may include task difficulty, level of project promotion, quality of researcher engagement with volunteers or even the inclusion of 'charismatic' features or species (consider penguins vs. wildebeest). Quantifying these variables and relating them to project measures represents a worthwhile future direction to this work. Although highly variable in the total number of classifications received, the classification curves of individual projects were highly comparable, with all but one project, *Supernova Hunters*, displaying a characteristic peak of activity upon launch followed by a rapid decline in classification activity. Supporting previous studies, when examining volunteer classification contribution inequality we found a large fraction of classifications are provided by a relatively small number of volunteers across all projects. Finally, demographic analysis of astronomy versus ecology projects indicated that projects appeal to a broad age-range regardless of academic domain. In contrast, examining gender differences revealed a clear male bias amongst astronomy projects, whereas ecology projects showed greater variability in their gender balance.

In addition to these central findings, we identified a number of exceptional features associated with a single project: *Supernova Hunters*. Beyond possessing a notably unusual classification curve that displayed weekly peaks of activity, reflecting the scheduled release of new data to this project concurrent with the sending of a newsletter to project volunteers, the *Supernova Hunters* project also displayed the highest level of classification contribution inequality amongst its volunteers. In the context of other findings here, it is also worth noting that the *Supernova Hunters* project displayed both an age and gender bias, with volunteers primarily males over the age of 65.

Our observation of frequent, weekly spikes of classification activity concurrent with the release of small sets of data in the *Supernova Hunters* project indicate that a scarcity of data is associated with sustained volunteer engagement with a project. For the *Supernova Hunters* project, this model of activity has arisen organically due to the incremental production of subject data, therefore, the pattern of volunteer engagement observed for this project is serendipitous rather than deliberately designed. Other researchers planning citizen science projects may wish to consider intentionally adopting a similar approach of incremental data release to encourage volunteer engagement. For example, a project with a large, pre-existing data set could partition their subjects for gradual release, generating an artificial scarcity to encourage more frequent volunteer interaction.

Adopting a model of generating artificial data scarcity within a project design may encourage increased volunteer interaction through creating a heightened sense of

competition to interact with a limited data set. A similar device of generating artificial scarcity is frequently applied in game or game-like design to drive further engagement [Deterding et al., 2011; Seaborn and Fels, 2015]. The motivation of competition is particularly relevant to a project such as *Supernova Hunters* where there is the legitimate possibility of being the first person to discover a supernova in the relatively small amount of new data released each week. It is understandable that engaged volunteers would be highly motivated to return to a project upon release of new data to have the opportunity of being the first to discover a supernova, or, as observed by Jackson et al., volunteers may just be motivated by being the first to see novel data [Jackson, Crowston et al., 2016].

However, our analyses indicate that perceived scarcity has effects beyond motivating volunteers to interact with a project more frequently. *Supernova Hunters* was identified as the most unequal project for volunteer classification contributions, therefore a small cohort of highly dedicated volunteers submit the majority of classifications in this project. The scarcity of data in this project, consequential competition to classify, and rapid processing of project subjects is causing the volunteer community of the *Supernova Hunters* project to be limited to a smaller group of highly dedicated individuals who are willing and able to return to the project upon data release.

There are a number of advantages to cultivating a community of dedicated and experienced volunteers who consistently return to a project upon data release, as this would enable quicker, and potentially more accurate, data processing. However, doing so may generate unexpected effects — our analyses indicate that this may be associated with a less diverse community of volunteers. We found the *Supernova Hunters* community to be demographically biased towards men of retirement age. It is possible that individuals who are unable to offer a regular, weekly time-commitment during the working day have been excluded from contributing to this project. Additionally, the increased competition to classify or higher science capital perceived to be required to contribute may make this project more appealing to older males, or, conversely, less appealing to other demographics. As shall be discussed, the impact of group diversity on project success remains to be fully delineated, particularly in online environments such as the Zooniverse where, although the majority of contributions are made independently and individually, there remains significant opportunity for group interaction and discovery through online fora [Boyajian et al., 2016]. Group diversity is commonly defined as differences in attributes between individuals resulting in the perception that others are different from oneself [van Knippenberg, De Dreu and Homan, 2004]. Although positive, negative or no relationships have been found between a group's diversity and its performance, the literature is in agreement that there are two primary mechanisms through which diversity can impact group performance; the 'informational' or 'decision-making perspective', or the 'social categorization perspective' [Chen, Ren and Riedl, 2010; van Knippenberg, De Dreu and Homan, 2004; van Knippenberg and Schippers, 2007; Williams and O'Reilly III, 1998].

As increased project participation has been found to develop the scientific knowledge of volunteers [Masters et al., 2016], cultivating the selection of a highly dedicated small community of volunteers through gamification may benefit a more challenging project that requires a higher level of skill, such as Foldit [Cooper et al.,

2010]. Although we do not examine the impact of sustained volunteer engagement on data quality in the *Supernova Hunters* project in the analyses presented here, this represents an interesting direction for future research. Further, sustained engagement may result in changed modes of volunteer participation over time, generating higher proportion of more experienced volunteers able to contribute to important community functions such as helping to onboard newcomers, answering questions and moderating discussions [Jackson, Østerlund et al., 2015; Mugar et al., 2014; Østerlund et al., 2014].

However, as we have seen here, cultivating a small, highly dedicated volunteer community can be associated with less diversity. As described in the social categorization perspective of the impact of group diversity on performance, smaller and more homogenous groups may outperform heterogeneous groups because people categorize themselves and others into social groups based on differences in social attributes, and as a result have a more positive experience when working with others they consider similar [Chen, Ren and Riedl, 2010]. Therefore, a more homogenous community has the potential to positively benefit a project. However, there are circumstances when a less-diverse community may have a negative impact on project success. For example, a cross-section of volunteers representative of the broader population may be essential for the validity of results in a health-related data collection project. Further, as stated in the 'informational' or 'decision-making perspective' of the impact of diversity on group performance, heterogeneous groups can outperform homogenous groups due to their broader range of skills, knowledge and opinions. Beyond research objectives, lack of group diversity may also curtail the potential of a citizen science project to achieve other aims, such as fostering scientific education within typically underserved communities. However, the effect of group diversity on the success of a citizen science project will be influenced by many factors, including the type and complexity of the task involved, task interdependency and the group type and size [Horwitz and Horwitz, 2007; van Knippenberg and Schippers, 2007]. Further, the importance of these variables will have differing importance dependent on whether a project is online or offline [Martins, Gilson and Maynard, 2004].

It should also be considered whether 'designing for exclusivity' is acceptable to the broader ethos of citizen science, and human-computer interaction more generally. Frequently, citizen science projects have the twin goals of contributing to both scientific productivity as well as 'social good', e.g. encouraging learning about and participation in science [Woodcock et al., 2017]. Those designing and implementing online citizen science projects should also consider whether it is acceptable to the practice of citizen science to cultivate scenarios that intentionally restrict the opportunity of the full volunteer community to access a project, or whether they should be implementing inclusive design approaches and 'designing for all'. Beyond the goal of encouraging diversity, researchers should consider the extent to which they can and should facilitate the accessibility of their project to underserved online communities, which may include people with low ICT skills, the elderly or individuals with reading difficulties, through the implementation of relevant inclusive design approaches such as universal usability [Shneiderman, 2000; Vanderheiden, 2000], user-sensitive inclusive design [Newell, 2011], designing for accessibility [Keates, 2006] and ability-based design [Wobbrock et al., 2011], and these factors should be examined closely in future work.

When considering the consequences of implementing design changes modelled on the observations made here, project owners must be cognisant of the dual aims of citizen science; to perform authentic research in the most efficient manner possible (scientific efficiency) and to allow a broad community to engage with real research (social inclusivity). The relationship between these two aims, social inclusivity and scientific efficiency, is nuanced. There may be instances where a project is more scientifically efficient if it is more exclusive, therefore, the scientific efficiency of a citizen science project may occasionally directly conflict with the aim of social inclusivity. Although designing projects for efficiency via exclusivity is not ideal, it is not clear whether the alternative, of reducing scientific efficiency in the name of inclusivity, is preferable. For example, in the *Supernova Hunters* project, social inclusivity may have been enhanced through providing greater opportunity to classify through increasing the number of classifications required to more than necessary for each image. However, this would not only represent a wasteful application of volunteer time and enthusiasm, but may undermine one of the most common motivations of volunteers — to make an authentic contribution [Raddick et al., 2010].

In conclusion, we present here quantitative evidence demonstrating how subtle changes to online citizen science project design can influence many facets of the nature of volunteer interaction, including who participates, when and how much. Through analysing the most comprehensive collection of citizen science project data gathered to date, we observe sustained volunteer engagement emerging from the incremental release of small amounts of data in a single Zooniverse project, *Supernova Hunters*. However, this increased interaction was observed in conjunction with high levels of classification contribution inequality, and a demographically biased community of volunteers. This model of incremental data release has cultivated a scenario where a large number of classifications are provided by a small community of volunteers, contrary to the typical Zooniverse project in which a small number of classifications are provided by a large community. These observations illustrate the tension that can exist between designing a citizen science project for scientific efficiency versus designing for social inclusivity, that stems from the liminal nature of citizen science between research and public engagement.

Table 6: Project characteristics. This table details multiple measurements associated with the 63 projects examined in this study, with all variables calculated for 100 days post-launch of each individual project. 'N_all_classifications': total number of classifications; 'N_all_volunteers': total number of unique, registered, volunteers; 'N_classifications_100_days': number of classifications made within 100 days post-project launch; 'Classifications_not_logged_in': number of classifications made by volunteers who weren't logged into the Zooniverse during 100 days post-project launch; 'Classifications_registered_users': number of classifications made to logged-in registered volunteers within 100 days post-project launch; 'N_unique_users': number of unique volunteers who classified during 100 days post-project launch; 'Least_classifications_single_user': least classified volunteer during 100 days post-project launch; 'Mean_classifications_per_user': mean classifications made by registered volunteers during 100 days post-project launch; 'SD_classifications_per_user': standard deviation of the classifications made by registered volunteers during 100 days post-project launch; 'Mode_classifications': mode classifications made by registered volunteers during 100 days post-project launch; 'Median_classifications': median classifications made by registered volunteers during 100 days post-project launch.

| Project_name | Domain | Launch_date | Launch_date_+100_days | N_all_classifications | N_all_volunteers | N_classifications_100_days | Classifications_not_logged_in | Classifications_registered_users | N_unique_users | Least_classifications_single_user | Most_classifications_single_user | Mean_classifications_per_user | SD_classifications_per_user | Mode_classifications | Median_classifications |
|------------------------------|---------------|-------------|-----------------------|-----------------------|------------------|----------------------------|-------------------------------|----------------------------------|----------------|-----------------------------------|----------------------------------|-------------------------------|-----------------------------|----------------------|------------------------|
| Microplants | Ecology | 30/06/2013 | 08/10/2013 | 42720 | 226 | 8577 | 1626 | 6951 | 17 | 2 | 4408 | 408.9 | 1115.0 | 2 | 6 |
| Microscopy Masters | Biomedical | 31/03/2016 | 09/07/2016 | 42070 | 2583 | 23404 | 9164 | 14240 | 1343 | 1 | 280 | 10.6 | 17.1 | 1 | 7 |
| Orchid Observers | Ecology | 23/04/2015 | 01/08/2015 | 53531 | 1769 | 23529 | 2712 | 20817 | 838 | 1 | 4012 | 24.8 | 175.5 | 1 | 3 |
| Planet Four: Craters | Astronomy | 31/03/2015 | 09/07/2015 | 32957 | 702 | 23702 | 12889 | 10813 | 546 | 1 | 226 | 19.8 | 32.5 | 1 | 7 |
| Arizona BatWatch | Ecology | 19/10/2016 | 27/01/2017 | 25702 | 1433 | 24389 | 5560 | 18829 | 1322 | 1 | 2646 | 14.2 | 86.0 | 1 | 4 |
| Season Spotter Image Marking | Ecology | 21/07/2015 | 29/10/2015 | 78849 | 3762 | 25102 | 9346 | 15756 | 1199 | 1 | 388 | 13.1 | 23.6 | 1 | 6 |
| Mapping Change | Ecology | 20/09/2016 | 29/12/2016 | 26769 | 1500 | 25218 | 2570 | 22648 | 1375 | 1 | 2197 | 16.5 | 89.7 | 1 | 3 |
| Shakespeare's World | Transcription | 08/12/2015 | 17/03/2016 | 96670 | 2153 | 41931 | 19517 | 22414 | 1147 | 1 | 1872 | 19.5 | 81.5 | 1 | 4 |
| Decoding the Civil War | Transcription | 21/06/2016 | 29/09/2016 | 70763 | 3104 | 45096 | 12740 | 32356 | 2300 | 1 | 1265 | 14.1 | 57.0 | 1 | 3 |
| AnnoTate | Transcription | 01/09/2015 | 10/12/2015 | 126321 | 1443 | 47088 | 31475 | 15613 | 724 | 1 | 2145 | 21.6 | 94.8 | 1 | 5 |
| Wildebeest Watch | Ecology | 30/06/2015 | 08/10/2015 | 109700 | 2846 | 50230 | 13378 | 36852 | 1399 | 1 | 2426 | 26.3 | 86.1 | 2 | 9 |

Continued on the next page

Table 6: Continued from the previous page.

| Project_name | Domain | Launch_date | Launch_date_+100_days | N_all_classifications | N_all_volunteers | N_classifications_100_days | Classifications_not_logged_in | Classifications_registered_users | N_unique_users | Least_classifications_single_user | Most_classifications_single_user | Mean_classifications_per_user | SD_classifications_per_user | Mode_classifications | Median_classifications |
|-------------------------------|---------------|-------------|-----------------------|-----------------------|------------------|----------------------------|-------------------------------|----------------------------------|----------------|-----------------------------------|----------------------------------|-------------------------------|-----------------------------|----------------------|------------------------|
| Stardate M83 | Astronomy | 13/01/2014 | 23/04/2014 | 52258 | 1193 | 52246 | 12693 | 39553 | 1193 | 1 | 5200 | 33.2 | 209.9 | 1 | 7 |
| Poppin' Galaxy | Astronomy | 23/03/2016 | 01/07/2016 | 111102 | 2601 | 54337 | 11590 | 42747 | 1255 | 1 | 3561 | 34.1 | 142.5 | 1 | 9 |
| Pattern Perception | Other | 23/05/2016 | 31/08/2016 | 80187 | 2409 | 63246 | 17755 | 45491 | 1726 | 1 | 2661 | 26.4 | 102.8 | 1 | 7 |
| Notes from Nature Launch 2016 | Transcription | 16/06/2016 | 24/09/2016 | 145534 | 2333 | 63640 | 5100 | 58540 | 1280 | 1 | 5733 | 45.7 | 275.5 | 1 | 5 |
| Season Spotter Questions | Ecology | 21/07/2015 | 29/10/2015 | 193992 | 4988 | 73247 | 25098 | 48149 | 2062 | 1 | 1372 | 23.4 | 61.9 | 1 | 10 |
| Whales as Individuals | Ecology | 30/06/2015 | 08/10/2015 | 271610 | 6902 | 95971 | 14804 | 81167 | 1861 | 1 | 4799 | 43.6 | 196.2 | 1 | 11 |
| Computer Vision: Serengeti | Ecology | 10/05/2016 | 18/08/2016 | 114647 | 3145 | 100248 | 17628 | 82620 | 2425 | 1 | 8394 | 34.1 | 200.7 | 1 | 10 |
| Understanding Animal Faces | Other | 21/07/2016 | 29/10/2016 | 136636 | 2290 | 110051 | 30627 | 79424 | 1999 | 1 | 3806 | 39.7 | 133.9 | 1 | 11 |
| Galaxy Zoo: Bar Lengths | Astronomy | 30/06/2015 | 08/10/2015 | 319528 | 5307 | 118236 | 17244 | 100992 | 1972 | 1 | 8642 | 51.2 | 302.2 | 1 | 11 |
| Radio Meteor Zoo | Astronomy | 12/08/2016 | 20/11/2016 | 159700 | 3194 | 122688 | 21010 | 101678 | 2769 | 1 | 4358 | 36.7 | 178.5 | 1 | 8 |
| Worm Watch Lab | Biomedical | 03/07/2013 | 11/10/2013 | 711111 | 13348 | 123236 | 34163 | 89073 | 4177 | 1 | 2901 | 21.3 | 75.8 | 2 | 8 |
| Cyclone Center | Other | 27/09/2012 | 05/01/2013 | 645686 | 13758 | 135832 | 36188 | 99644 | 3016 | 1 | 9987 | 33.0 | 215.7 | 6 | 9 |
| Comet Hunters | Astronomy | 16/12/2015 | 25/03/2016 | 760203 | 7157 | 138741 | 7020 | 131721 | 1458 | 1 | 7062 | 90.3 | 403.7 | 1 | 14 |

Continued on the next page

Table 6: Continued from the previous page.

| Project_name | Domain | Launch_date | Launch_date_+100_days | N_all_classifications | N_all_volunteers | N_classifications_100_days | Classifications_not_logged_in | Classifications_registered_users | N_unique_users | Least_classifications_single_user | Most_classifications_single_user | Mean_classifications_per_user | SD_classifications_per_user | Mode_classifications | Median_classifications |
|------------------------------------|---------------|-------------|-----------------------|-----------------------|------------------|----------------------------|-------------------------------|----------------------------------|----------------|-----------------------------------|----------------------------------|-------------------------------|-----------------------------|----------------------|------------------------|
| Jungle Rhythms | Ecology | 13/12/2015 | 22/03/2016 | 280240 | 2871 | 141972 | 16629 | 125343 | 1278 | 1 | 15311 | 98.1 | 699.5 | 1 | 12 |
| Bat Detective | Ecology | 02/10/2012 | 10/01/2013 | 587871 | 5501 | 174490 | 31600 | 142890 | 1183 | 1 | 20036 | 120.8 | 844.0 | 1 | 10 |
| Planet Four: Terrains | Astronomy | 30/06/2015 | 08/10/2015 | 905464 | 10097 | 186343 | 44796 | 141547 | 2481 | 1 | 8636 | 57.1 | 324.9 | 1 | 15 |
| Snapshots at Sea | Ecology | 22/12/2015 | 31/03/2016 | 964128 | 4575 | 193842 | 24230 | 169612 | 1174 | 1 | 21685 | 144.5 | 949.2 | 4 | 24 |
| Condor Watch | Ecology | 15/04/2014 | 24/07/2014 | 484759 | 5796 | 210787 | 66617 | 144170 | 3100 | 1 | 5746 | 46.5 | 217.5 | 1 | 11 |
| Science Gossip | Transcription | 04/03/2015 | 12/06/2015 | 538470 | 7850 | 216012 | 39981 | 176031 | 4321 | 1 | 11394 | 40.7 | 216.2 | 1 | 14 |
| Notes from Nature Launch 2013 | Transcription | 23/04/2013 | 01/08/2013 | 1011392 | 9133 | 229983 | 78105 | 151878 | 3354 | 1 | 4802 | 45.3 | 206.5 | 1 | 7 |
| Western Shield — Camera Watch | Ecology | 07/04/2016 | 16/07/2016 | 472255 | 3367 | 237321 | 24070 | 213251 | 1718 | 1 | 10178 | 124.1 | 551.6 | 1 | 16 |
| Chicago Wildlife Watch Launch 2014 | Ecology | 11/09/2014 | 20/12/2014 | 2994189 | 8286 | 267644 | 70386 | 197258 | 805 | 1 | 20187 | 245.0 | 999.5 | 3 | 32 |
| Chicago Wildlife Watch Launch 2016 | Ecology | 26/10/2016 | 03/02/2017 | 320955 | 2121 | 299342 | 42174 | 257168 | 1887 | 1 | 7163 | 136.3 | 374.7 | 1 | 34 |
| Operation War Diary | Transcription | 14/01/2014 | 24/04/2014 | 795960 | 14836 | 316246 | 43873 | 272373 | 8523 | 1 | 6472 | 32.0 | 159.7 | 1 | 5 |
| Sunspotter | Astronomy | 27/02/2014 | 07/06/2014 | 5604299 | 8578 | 324311 | 81516 | 242795 | 548 | 1 | 34510 | 443.1 | 2019.8 | 5 | 92 |

Continued on the next page

Table 6: Continued from the previous page.

| Project_name | Domain | Launch_date | Launch_date_+100_days | N_all_classifications | N_all_volunteers | N_classifications_100_days | Classifications_not_logged_in | Classifications_registered_users | N_unique_users | Least_classifications_single_user | Most_classifications_single_user | Mean_classifications_per_user | SD_classifications_per_user | Mode_classifications | Median_classifications |
|-------------------------------|------------|-------------|-----------------------|-----------------------|------------------|----------------------------|-------------------------------|----------------------------------|----------------|-----------------------------------|----------------------------------|-------------------------------|-----------------------------|----------------------|------------------------|
| Plankton Portal | Ecology | 17/09/2013 | 26/12/2013 | 1378040 | 11607 | 341568 | 75560 | 266008 | 3331 | 1 | 51345 | 79.9 | 1211.5 | 1 | 8 |
| Fossil Finder | Other | 15/08/2015 | 23/11/2015 | 681833 | 6499 | 383920 | 88358 | 295562 | 2730 | 1 | 38717 | 108.3 | 935.0 | 1 | 7 |
| Snapshot Wisconsin | Ecology | 17/05/2016 | 25/08/2016 | 730164 | 3489 | 397765 | 83724 | 314041 | 2566 | 1 | 8381 | 122.4 | 374.4 | 1 | 26 |
| Wisconsin Wildlife Watch | Ecology | 08/01/2016 | 17/04/2016 | 751528 | 3641 | 561024 | 53264 | 507760 | 2572 | 1 | 29181 | 197.4 | 828.2 | 1 | 37 |
| Cell Slider | Biomedical | 24/10/2012 | 01/02/2013 | 2757067 | 25688 | 565328 | 284985 | 280343 | 5110 | 1 | 13051 | 54.9 | 328.8 | 5 | 14 |
| Milky Way Project Launch 2016 | Astronomy | 15/09/2016 | 24/12/2016 | 629595 | 2992 | 570611 | 36206 | 534405 | 2821 | 1 | 32577 | 189.4 | 1494.4 | 1 | 17 |
| Radio Galaxy Zoo | Astronomy | 17/12/2013 | 27/03/2014 | 1956630 | 12179 | 597090 | 166055 | 431035 | 4173 | 1 | 35564 | 103.3 | 770.1 | 1 | 13 |
| Disk Detective | Astronomy | 30/01/2014 | 10/05/2014 | 2222273 | 12122 | 616142 | 198983 | 417159 | 4600 | 1 | 32183 | 90.7 | 981.0 | 1 | 11 |
| Higgs Hunters | Other | 26/11/2014 | 06/03/2015 | 1375450 | 10643 | 654836 | 96728 | 558108 | 5239 | 1 | 18013 | 106.5 | 585.5 | 1 | 18 |
| Supernova Hunters | Astronomy | 12/07/2016 | 20/10/2016 | 1121725 | 2866 | 716055 | 18500 | 697555 | 1689 | 1 | 103069 | 413.0 | 3722.1 | 1 | 17 |
| Penguin Watch | Ecology | 17/09/2014 | 26/12/2014 | 4660524 | 39145 | 824254 | 163634 | 660620 | 9117 | 1 | 10059 | 72.5 | 321.7 | 1 | 16 |
| Gravity Spy | Astronomy | 12/10/2016 | 20/01/2017 | 1014226 | 4629 | 872439 | 83287 | 789152 | 3854 | 1 | 20160 | 204.8 | 758.0 | 10 | 40 |
| Milky Way Project Launch 2013 | Astronomy | 12/12/2013 | 22/03/2014 | 2192295 | 22530 | 944727 | 204365 | 740362 | 9640 | 1 | 20459 | 76.8 | 449.2 | 1 | 7 |
| Floating Forests | Ecology | 07/08/2014 | 15/11/2014 | 2782903 | 6818 | 1032143 | 320568 | 711575 | 2810 | 1 | 79229 | 253.2 | 2211.8 | 1 | 36 |

Continued on the next page

Table 6: Continued from the previous page.

| Project_name | Domain | Launch_date | Launch_date_+100_days | N_all_classifications | N_all_volunteers | N_classifications_100_days | Classifications_not_logged_in | Classifications_registered_users | N_unique_users | Least_classifications_single_user | Most_classifications_single_user | Mean_classifications_per_user | SD_classifications_per_user | Mode_classifications | Median_classifications |
|----------------------------|-----------|-------------|-----------------------|-----------------------|------------------|----------------------------|-------------------------------|----------------------------------|----------------|-----------------------------------|----------------------------------|-------------------------------|-----------------------------|----------------------|------------------------|
| Chimp & See | Ecology | 22/04/2015 | 31/07/2015 | 2815482 | 7438 | 1076635 | 392514 | 684121 | 3862 | 1 | 28539 | 177.1 | 706.9 | 1 | 32 |
| Andromeda Project | Astronomy | 05/12/2012 | 15/03/2013 | 1958717 | 11255 | 1078393 | 153592 | 924801 | 5302 | 1 | 11183 | 174.4 | 550.6 | 1 | 27 |
| Planet Hunters Launch 2014 | Astronomy | 18/09/2014 | 27/12/2014 | 8829803 | 43264 | 1085355 | 124803 | 960552 | 6017 | 1 | 47715 | 159.6 | 1068.5 | 1 | 23 |
| Seafloor Explorer | Ecology | 13/09/2012 | 22/12/2012 | 2662916 | 25356 | 1159534 | 174013 | 985521 | 9147 | 1 | 10586 | 107.7 | 373.9 | 3 | 23 |
| Asteroid Zoo | Astronomy | 24/06/2014 | 02/10/2014 | 2664806 | 12322 | 1209318 | 375562 | 833756 | 5571 | 1 | 35678 | 149.7 | 726.7 | 1 | 23 |
| Snapshot Supernova | Astronomy | 19/03/2015 | 27/06/2015 | 2021281 | 4906 | 1461538 | 905348 | 556190 | 3871 | 1 | 44609 | 143.7 | 883.0 | 1 | 36 |
| WildCam Gorongosa | Ecology | 10/09/2015 | 19/12/2015 | 4085550 | 15663 | 1797998 | 610319 | 1187679 | 4657 | 1 | 26477 | 255.0 | 1218.7 | 1 | 30 |
| Camera Catalogue | Ecology | 04/08/2016 | 12/11/2016 | 2563684 | 5854 | 1855799 | 162096 | 1693703 | 4759 | 1 | 40619 | 355.9 | 1321.3 | 1 | 48 |
| Pulsar Hunters | Astronomy | 12/01/2016 | 21/04/2016 | 3210326 | 10643 | 3049023 | 323717 | 2725306 | 10272 | 1 | 9318 | 265.3 | 493.0 | 1 | 100 |
| Galaxy Zoo — CAN-DELS | Astronomy | 11/09/2012 | 20/12/2012 | 25335564 | 153129 | 3699564 | 572620 | 3126944 | 21353 | 1 | 58637 | 146.4 | 927.0 | 1 | 22 |
| Planet Four | Astronomy | 08/01/2013 | 18/04/2013 | 5263065 | 59534 | 3742864 | 1116670 | 2626194 | 34936 | 1 | 37461 | 75.2 | 466.7 | 2 | 17 |
| Snapshot Serengeti | Ecology | 11/12/2012 | 21/03/2013 | 22634263 | 50968 | 6402357 | 903166 | 5499191 | 18543 | 1 | 22378 | 296.6 | 862.5 | 1 | 58 |
| Space Warps | Astronomy | 08/05/2013 | 16/08/2013 | 18596661 | 30116 | 8011730 | 492064 | 7519666 | 9631 | 1 | 250190 | 780.8 | 6271.0 | 1 | 63 |

References

- Boyajian, T. S., LaCourse, D. M., Rappaport, S. A., Fabrycky, D., Fischer, D. A., Gandolfi, D., Kennedy, G. M., Korhonen, H., Liu, M. C., Moor, A., Olah, K., Vida, K., Wyatt, M. C., Best, W. M. J., Brewer, J., Ciesla, F., Csák, B., Deeg, H. J., Dupuy, T. J., Handler, G., Heng, K., Howell, S. B., Ishikawa, S. T., Kovács, J., Kozakis, T., Kriskovics, L., Lehtinen, J., Lintott, C., Lynn, S., Nespral, D., Nikbakhsh, S., Schawinski, K., Schmitt, J. R., Smith, A. M., Szabo, G., Szabo, R., Viuhho, J., Wang, J., Weiksnar, A., Bosch, M., Connors, J. L., Goodman, S., Green, G., Hoekstra, A. J., Jebson, T., Jek, K. J., Omohundro, M. R., Schwengeler, H. M. and Szewczyk, A. (2016). 'Planet Hunters IX. KIC 8462852 — where's the flux?' *Monthly Notices of the Royal Astronomical Society* 457 (4), pp. 3988–4004. <https://doi.org/10.1093/mnras/stw218>.
- Brabham, D. C. (2008). 'Crowdsourcing as a model for problem solving'. *Convergence: The International Journal of Research into New Media Technologies* 14 (1), pp. 75–90. <https://doi.org/10.1177/1354856507084420>.
- Candido dos Reis, F. J., Lynn, S., Ali, H. R., Eccles, D., Hanby, A., Provenzano, E., Caldas, C., Howat, W. J., McDuffus, L.-A., Liu, B., Daley, F., Coulson, P., Vyas, R. J., Harris, L. M., Owens, J. M., Carton, A. F. M., McQuillan, J. P., Paterson, A. M., Hirji, Z., Christie, S. K., Holmes, A. R., Schmidt, M. K., Garcia-Closas, M., Easton, D. F., Bolla, M. K., Wang, Q., Benitez, J., Milne, R. L., Mannermaa, A., Couch, F., Devilee, P., Tollenaar, R. A. E. M., Seynaeve, C., Cox, A., Cross, S. S., Blows, F. M., Sanders, J., Groot, R. de, Figueroa, J., Sherman, M., Hooning, M., Brenner, H., Holleczeck, B., Stegmaier, C., Lintott, C. and Pharoah, P. D. P. (2015). 'Crowdsourcing the general public for large scale molecular pathology studies in cancer'. *EBioMedicine* 2 (7), pp. 681–689. <https://doi.org/10.1016/j.ebiom.2015.05.009>.
- Cardamone, C., Schawinski, K., Sarzi, M., Bamford, S. P., Bennert, N., Urry, C. M., Lintott, C., Keel, W. C., Parejko, J., Nichol, R. C., Thomas, D., Andreescu, D., Murray, P., Raddick, M. J., Slosar, A., Szalay, A. and VandenBerg, J. (2009). 'Galaxy zoo green peas: discovery of a class of compact extremely star-forming galaxies'. *Monthly Notices of the Royal Astronomical Society* 399 (3), pp. 1191–1205. <https://doi.org/10.1111/j.1365-2966.2009.15383.x>.
- Chen, J., Ren, Y. and Riedl, J. (2010). 'The effects of diversity on group productivity and member withdrawal in online volunteer groups'. In: *Proceedings of the 28th international conference on Human factors in computing systems — CHI '10*. ACM Press. <https://doi.org/10.1145/1753326.1753447>.
- Clark, D. J., Nicholas, D. and Jamali, H. R. (2014). 'Evaluating information seeking and use in the changing virtual world: the emerging role of Google Analytics'. *Learned Publishing* 27 (3), pp. 185–194. <https://doi.org/10.1087/20140304>.
- Cooper, S., Khatib, F., Treuille, A., Barbero, J., Lee, J., Beenen, M., Leaver-Fay, A., Baker, D., Popović, Z. and Players, F. (2010). 'Predicting protein structures with a multiplayer online game'. *Nature* 466 (7307), pp. 756–760. <https://doi.org/10.1038/nature09304>.
- Cox, J., Oh, E. Y., Simmons, B., Lintott, C. L., Masters, K. L., Greenhill, A., Graham, G. and Holmes, K. (2015). 'Defining and Measuring Success in Online Citizen Science: A Case Study of Zooniverse Projects'. *Computing in Science & Engineering* 17 (4), pp. 28–41. <https://doi.org/10.1109/MCSE.2015.65>.
- Crutzen, R., Roosjen, J. L. and Poelman, J. (2013). 'Using Google Analytics as a process evaluation method for Internet-delivered interventions: an example on sexual health'. *Health Promotion International* 28 (1), pp. 36–42. <https://doi.org/10.1093/heapro/das008>.

- Deterding, S., Dixon, D., Khaled, R. and Nacke, L. (2011). 'From game design elements to gamefulness: defining "gamification"'. In: *Proceedings of the 15th International Academic MindTrek Conference on Envisioning Future Media Environments — MindTrek '11* (Tampere, Finland). New York, U.S.A.: ACM Press, pp. 9–15. <https://doi.org/10.1145/2181037.2181040>.
- Forrester, T. D., Baker, M., Costello, R., Kays, R., Parsons, A. W. and McShea, W. J. (2017). 'Creating advocates for mammal conservation through citizen science'. *Biological Conservation* 208, pp. 98–105. <https://doi.org/10.1016/j.biocon.2016.06.025>.
- Franke, N. and von Hippel, E. (2003). 'Satisfying heterogeneous user needs via innovation toolkits: the case of Apache security software'. *Research Policy* 32 (7), pp. 1199–1215. [https://doi.org/10.1016/s0048-7333\(03\)00049-0](https://doi.org/10.1016/s0048-7333(03)00049-0).
- Gastwirth, J. L. (1972). 'The estimation of the Lorenz curve and Gini index'. *The Review of Economics and Statistics* 54 (3), p. 306. <https://doi.org/10.2307/1937992>.
- Google (2018). *About demographics and interests*. URL: <https://support.google.com/analytics/answer/2799357?hl=en> (visited on 2018).
- Horwitz, S. K. and Horwitz, I. B. (2007). 'The effects of team diversity on team outcomes: a meta-analytic review of team demography'. *Journal of Management* 33 (6), pp. 987–1015. <https://doi.org/10.1177/0149206307308587>.
- Jackson, C. B., Crowston, K., Mugar, G. and Østerlund, C. (2016). "'Guess what! You're the first to see this event'". In: *Proceedings of the 19th International Conference on Supporting Group Work — GROUP '16* (Sanibel Island, FL, U.S.A.). ACM, pp. 171–179. <https://doi.org/10.1145/2957276.2957284>.
- Jackson, C. B., Østerlund, C., Mugar, G., Hassman, K. D. and Crowston, K. (2015). 'Motivations for sustained participation in crowdsourcing: case studies of citizen science on the role of talk'. In: *2015 48th Hawaii International Conference on System Sciences*. IEEE, pp. 1624–1634. <https://doi.org/10.1109/hicss.2015.196>.
- Jeppesen, L. B. and Lakhani, K. R. (2010). 'Marginality and problem-solving effectiveness in broadcast search'. *Organization Science* 21 (5), pp. 1016–1033. <https://doi.org/10.1287/orsc.1090.0491>.
- Kawrykow, A., Roumanis, G., Kam, A., Kwak, D., Leung, C., Wu, C., Zarour, E., Sarmenta, L., Blanchette, M. and and, J. W. (2012). 'Phylo: a citizen science approach for improving multiple sequence alignment'. *PLoS ONE* 7 (3), e31362. <https://doi.org/10.1371/journal.pone.0031362>.
- Keates, S. (2006). *Designing for accessibility: a business guide to countering design exclusion*. Boca Raton, FL, U.S.A.: CRC Press. <https://doi.org/10.1201/9781482269543>.
- Khatib, F., DiMaio, F., Cooper, S., Kazmierczyk, M., Gilski, M., Krzywda, S., Zabranska, H., Pichova, I., Thompson, J., Popović, Z., Jaskolski, M. and Baker, D. (2011). 'Crystal structure of a monomeric retroviral protease solved by protein folding game players'. *Nature Structural & Molecular Biology* 18 (10), pp. 1175–1177. <https://doi.org/10.1038/nsmb.2119>.
- Kim, J. S., Greene, M. J., Zlateski, A., Lee, K., Richardson, M., Turaga, S. C., Purcaro, M., Balkam, M., Robinson, A., Behabadi, B. F., Campos, M., Denk, W. and Seung, H. S. (2014). 'Space-time wiring specificity supports direction selectivity in the retina'. *Nature* 509 (7500), pp. 331–336. <https://doi.org/10.1038/nature13240>.

- Lee, T. K., Crowston, K., Østerlund, C. and Miller, G. (2017). 'Recruiting messages matter: message strategies to attract citizen scientists'. In: *Companion of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing — CSCW '17 Companion* (Portland, OR, U.S.A.). ACM Press, pp. 227–230.
<https://doi.org/10.1145/3022198.3026335>.
- Lintott, C. J., Schawinski, K., Slosar, A., Land, K., Bamford, S., Thomas, D., Raddick, M. J., Nichol, R. C., Szalay, A., Andreescu, D., Murray, P. and Vandenberg, J. (2008). 'Galaxy Zoo: morphologies derived from visual inspection of galaxies from the Sloan Digital Sky Survey'. *Monthly Notices of the Royal Astronomical Society* 389 (3), pp. 1179–1189.
<https://doi.org/10.1111/j.1365-2966.2008.13689.x>.
- Martins, L. L., Gilson, L. L. and Maynard, M. T. (2004). 'Virtual teams: what do we know and where do we go from here?' *Journal of Management* 30 (6), pp. 805–835. <https://doi.org/10.1016/j.jm.2004.05.002>.
- Masters, K., Oh, E. Y., Cox, J., Simmons, B., Lintott, C., Graham, G., Greenhill, A. and Holmes, K. (2016). 'Science learning via participation in online citizen science'. *JCOM* 15 (3), A07.
 URL: https://jcom.sissa.it/archive/15/03/JCOM_1503_2016_A07.
- Mugar, G., Osterlund, C., DeVries Hassman, K., Crowston, K. and Jackson, C. B. (2014). 'Planet Hunters and Seafloor Explorers: Legitimate Peripheral Participation Through Practice Proxies in Online Citizen Science'. Paper presented at the Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing, Baltimore, Maryland, U.S.A. In: <https://doi.org/10.1145/2531602.2531721>.
- Newell, A. F. (2011). 'Design and the digital divide: insights from 40 years in computer support for older and disabled people'. *Synthesis Lectures on Assistive, Rehabilitative and Health-Preserving Technologies* 1 (1), pp. 1–195.
<https://doi.org/10.2200/s00369ed1v01y201106arh001>.
- Ortega, F., Gonzalez-Barahona, J. M. and Robles, G. (2008). 'On the inequality of contributions to Wikipedia'. In: *Proceedings of the 41st Annual Hawaii International Conference on System Sciences (HICSS 2008)*. IEEE.
<https://doi.org/10.1109/hicss.2008.333>.
- Østerlund, C. S., Mugar, G., Jackson, C., Hassman, K. D. and Crowston, K. (2014). 'Socializing the crowd: learning to talk in citizen science'. *Academy of Management Proceedings* 2014 (1), p. 16799.
<https://doi.org/10.5465/ambpp.2014.16799abstract>.
- Pancierera, K., Halfaker, A. and Terveen, L. (2009). 'Wikipedians are born, not made: a study of power editors on Wikipedia'. In: *Proceedings of the ACM 2009 international conference on Supporting group work — GROUP '09*. ACM Press, pp. 51–60. <https://doi.org/10.1145/1531674.1531682>.
- Pimm, S. L., Jenkins, C. N., Abell, R., Brooks, T. M., Gittleman, J. L., Joppa, L. N., Raven, P. H., Roberts, C. M. and Sexton, J. O. (2014). 'The biodiversity of species and their rates of extinction, distribution and protection'. *Science* 344 (6187), pp. 1246752–1246752. <https://doi.org/10.1126/science.1246752>.
- Plaza, B. (2009). 'Monitoring web traffic source effectiveness with Google Analytics: an experiment with time series'. *Aslib Proceedings* 61 (5), pp. 474–482.
<https://doi.org/10.1108/00012530910989625>.
- (2011). 'Google Analytics for measuring website performance'. *Tourism Management* 32 (3), pp. 477–481.
<https://doi.org/10.1016/j.tourman.2010.03.015>.

- Raddick, M. J., Bracey, G., Gay, P. L., Lintott, C. J., Murray, P., Schawinski, K., Szalay, A. S. and Vandenberg, J. (2010). 'Galaxy Zoo: Exploring the Motivations of Citizen Science Volunteers'. *Astronomy Education Review* 9 (1), 010103, pp. 1–18. <https://doi.org/10.3847/AER2009036>. arXiv: 0909.2925.
- Sauermann, H. and Franzoni, C. (2015). 'Crowd science user contribution patterns and their implications'. *Proceedings of the National Academy of Sciences* 112 (3), pp. 679–684. <https://doi.org/10.1073/pnas.1408907112>.
- Seaborn, K. and Fels, D. I. (2015). 'Gamification in theory and action: a survey'. *International Journal of Human-Computer Studies* 74, pp. 14–31. <https://doi.org/10.1016/j.ijhcs.2014.09.006>.
- Segal, A., Gal, Y., Kamar, E., Horvitz, E., Bowyer, A. and Miller, G. (2016). 'Intervention strategies for increasing engagement in crowdsourcing: platform, predictions and experiments'. In: *Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence — IJCAI'16*. New York, NY, U.S.A.: AAAI Press, pp. 3861–3867. URL: <http://dl.acm.org/citation.cfm?id=3061053.3061159>.
- Shneiderman, B. (2000). 'Universal usability'. *Communications of the ACM* 43 (5), pp. 84–91. <https://doi.org/10.1145/332833.332843>.
- Simpson, R., Page, K. R. and De Roure, D. (2014). 'Zooniverse: observing the World's Largest Citizen Science Platform'. In: *Proceedings of the 23rd International Conference on World Wide Web — WWW '14 Companion* (Seoul, Korea, 7th–11th April 2014). ACM Press, pp. 1049–1054. <https://doi.org/10.1145/2567948.2579215>.
- Song, M. J., Ward, J., Choi, F., Nikoo, M., Frank, A., Shams, F., Tabi, K., Vigo, D. and Krausz, M. (2018). 'A process evaluation of a web-based mental health portal (WalkAlong) using Google Analytics'. *JMIR Mental Health* 5 (3), e50. <https://doi.org/10.2196/mental.8594>.
- Spiers, H., Swanson, A., Fortson, L., Simmons, B. D., Trouille, L., Blickhan, S. and Lintott, C. (2018). 'Patterns of volunteer behaviour across online citizen science'. In: *Companion of the The Web Conference 2018 on The Web Conference 2018 — WWW '18* (Lyon, France). <https://doi.org/10.1145/3184558.3186945>.
- Sprinks, J., Wardlaw, J., Houghton, R., Bamford, S. and Morley, J. (2017). 'Task workflow design and its impact on performance and volunteers' subjective preference in Virtual Citizen Science'. *International Journal of Human-Computer Studies* 104, pp. 50–63. <https://doi.org/10.1016/j.ijhcs.2017.03.003>.
- Straub, M. C. P. (2016). 'Giving citizen scientists a chance: a study of volunteer-led scientific discovery'. *Citizen Science: Theory and Practice* 1 (1), p. 5. <https://doi.org/10.5334/cstp.40>.
- Swanson, A., Arnold, T., Kosmala, M., Forester, J. and Packer, C. (2016). 'In the absence of a "landscape of fear": how lions, hyenas and cheetahs coexist'. *Ecology and Evolution* 6 (23), pp. 8534–8545. <https://doi.org/10.1002/ece3.2569>.
- Tsueng, G., Nanis, S. M., Fouquier, J., Good, B. M. and Su, A. I. (2016). 'Citizen science for mining the biomedical literature'. *Citizen Science: Theory and Practice* 1 (2), p. 14. <https://doi.org/10.5334/cstp.56>.
- Turner, S. J. (2010). 'Website statistics 2.0: using Google Analytics to measure library website effectiveness'. *Technical Services Quarterly* 27 (3), pp. 261–278. <https://doi.org/10.1080/07317131003765910>.

- van Knippenberg, D., De Dreu, C. K. W. and Homan, A. C. (2004). 'Work group diversity and group performance: an integrative model and research agenda'. *Journal of Applied Psychology* 89 (6), pp. 1008–1022.
<https://doi.org/10.1037/0021-9010.89.6.1008>.
- van Knippenberg, D. and Schippers, M. C. (2007). 'Work group diversity'. *Annual Review of Psychology* 58 (1), pp. 515–541.
<https://doi.org/10.1146/annurev.psych.58.110405.085546>.
- Vanderheiden, G. (2000). 'Fundamental principles and priority setting for universal usability'. In: *Proceedings on the 2000 conference on Universal Usability – CUU '00*. ACM Press. <https://doi.org/10.1145/355460.355469>.
- Wilkinson, D. M. (2008). 'Strong regularities in online peer production'. In: *Proceedings of the 9th ACM conference on Electronic commerce — EC '08* (Chicago, IL, U.S.A.). ACM Press, pp. 302–309.
<https://doi.org/10.1145/1386790.1386837>.
- Williams, K. Y. and O'Reilly III, C. A. (1998). 'Demography and diversity in organizations: a review of 40 years of research'. *Research in organizational behavior* 20, pp. 77–140.
- Wobbrock, J. O., Kane, S. K., Gajos, K. Z., Harada, S. and Froehlich, J. (2011). 'Ability-based design: concept, principles and examples'. *ACM Transactions on Accessible Computing* 3 (3), pp. 1–27.
<https://doi.org/10.1145/1952383.1952384>.
- Woodcock, J., Greenhill, A., Holmes, K., Graham, G., Cox, J., Oh, E. Y. and Masters, K. (2017). 'Crowdsourcing citizen science: exploring the tensions between paid professionals and users'. *Journal of Peer Production*.
 URL: <http://peerproduction.net/issues/issue-10-peer-production-and-work/peer-reviewed-papers/crowdsourcing-citizen-science-exploring-the-tensions-between-paid-professionals-and-users/>.
- Wright, D. E., Lintott, C. J., Smartt, S. J., Smith, K. W., Fortson, L., Trouille, L., Allen, C. R., Beck, M., Bouslog, M. C., Boyer, A., Chambers, K. C., Flewelling, H., Granger, W., Magnier, E. A., McMaster, A., Miller, G. R. M., O'Donnell, J. E., Simmons, B., Spiers, H., Tonry, J. L., Veldthuis, M., Wainscoat, R. J., Waters, C., Willman, M., Wolfenbarger, Z. and Young, D. R. (2017). 'A transient search using combined human and machine classifications'. *Monthly Notices of the Royal Astronomical Society* 472 (2), pp. 1315–1323.
<https://doi.org/10.1093/mnras/stx1812>.

Authors

Helen Spiers is the Biomedical Research Lead for The Zooniverse citizen science platform and a Postdoctoral Research Associate based in the Department of Astrophysics at the University of Oxford. Prior to starting her current position she obtained a PhD in developmental epigenetics from King's College London.
 E-mail: helen.spiers@physics.ox.ac.uk.

Alexandra Swanson is a behavioral and community ecologist interested in finding innovative technological solutions to conservation issues. As part of her PhD on mechanisms of coexistence among carnivores, she established a large-scale camera trap survey and partnered with the Zooniverse to build a citizen science platform to engage volunteers in classifying the millions of photographs produced. She formerly led the ecology initiatives at the Zooniverse and is now a AAAS Science and Technology Policy Fellow working on wildlife conservation policy issues.
 E-mail: ali@zooniverse.org.

Lucy Fortson is a Professor of Physics at the University of Minnesota. She is an observational astrophysicist studying galaxies and a cofounder of the Zooniverse citizen science platform. In collaboration with the University of Oxford and Adler Planetarium, Fortson leads the Zooniverse team developing human-computation algorithms to maximize the utility of the citizen science method of data processing. These methods are critical to extracting knowledge from the large amounts of data produced across the disciplines from astronomy to zoology. E-mail: lfortson@umn.edu.

Brooke Simmons is a lecturer in Physics at Lancaster University. She is Deputy Project Scientist for Galaxy Zoo and leads the Zooniverse's humanitarian and disaster relief arm. Dr Simmons' research uses citizen science for pattern detection in multiple disciplines, including astrophysics. E-mail: b.simmons@lancaster.ac.uk.

Laura Trouille is Vice President of Citizen Science at the Adler Planetarium in Chicago, co-Principal Investigator for the Zooniverse, and a Research Associate at Northwestern University. While earning her Ph.D. in astrophysics, she also earned the Center for the Integration of Research, Teaching and Learning's Delta certificate for STEM education research. The Adler-Zooniverse group collaborates closely with the Zooniverse web developers and team of researchers at the University of Oxford and the University of Minnesota. E-mail: trouille@zooniverse.org.

Samantha Blickhan is the Humanities Lead for the Zooniverse, and IMLS Postdoctoral Fellow at the Adler Planetarium in Chicago, IL. She received her PhD in musicology from Royal Holloway, University of London. E-mail: samantha@zooniverse.org.

Chris Lintott is a professor of astrophysics at the University of Oxford, where he is also a research fellow at New College. He is the principal investigator of both the Zooniverse and Galaxy Zoo, and his work includes investigations of galaxy evolution, planet populations and the possibilities of human-machine collaboration. E-mail: chris.lintott@physics.ox.ac.uk.

How to cite

Spiers, H., Swanson, A., Fortson, L., Simmons, B. D., Trouille, L., Blickhan, S. and Lintott, C. (2019). 'Everyone counts? Design considerations in online citizen science'. *JCOM* 18 (01), A04. <https://doi.org/10.22323/2.18010204>.



© The Author(s). This article is licensed under the terms of the Creative Commons Attribution — NonCommercial — NoDerivatives 4.0 License. ISSN 1824-2049. Published by SISSA Medialab. jcom.sissa.it