Cox J., Oh E.Y., Simmons B., Graham G., Greenhill A., Lintott C., Masters K., Woodcock J.

Doing good online: the changing relationships between motivations, activity and retention among online volunteers

Article (Accepted version) (Refereed)

Original citation: Cox, Joe, Oh, Eun Young, Simmons, B., Graham, G., Greenhill, A., Lintott, C., Masters, K. and Woodcock, Jamie (2017) *Doing good online: the changing relationships between motivations, activity and retention among online volunteers.* Nonprofit and Voluntary Sector Quarterly . ISSN 0899-7640

Doing good online: The changing relationships between motivations, activity and retention among online volunteers

Dr Joe Cox; University of Portsmouth Business School, Richmond Building, Portland Street, Portsmouth, PO1 3DE, UK; +44 2392 844723; joe.cox@port.ac.uk

Dr. Joe Cox is a Principal Lecturer at the University of Portsmouth Business School. His research interests relate to the digital economy and areas such as crowdsourcing, crowdfunding, digital piracy and video games. He is also lead academic investigator on research projects funded by the EPSRC and Innovate UK.

Dr Eun Young Oh; University of Portsmouth Business School, Richmond Building, Portland Street, Portsmouth, PO1 3DE, UK; +44 2392 844159; ey.oh@port.ac.uk

Dr. Eun Young Oh is a senior research associate at the University of Portsmouth Business School. She has developed an expertise in economics over several years of industrial and academic experience. Now Dr. Oh has transferred the skills she honed studying Monetary Policy, DSGE Models, Bayesian Analysis and Panel Data Analyses to the Digital Economy, a new and exciting field of research.

Dr Brooke Simmons; University of California, San Diego, Center for Astrophysics and Space Sciences, 9500 Gilman Drive #0424, CA La Jolla 92093, USA; +1 858 822 4463; bdsimmons@ucsd.edu

Dr. Brooke Simmons is an astrophysicist researching black holes and galaxies at the University of California, San Diego. Dr. Simmons is a member of the Zooniverse team, involved in planning and building several different projects within the Zooniverse, as well as interacting with the large community of volunteers.

Dr Gary Graham; Leeds University Business School, Maurice Keyworth Building, leeds, West Yorkshire, LS9 9LB, UK; +44 113 343 8557, g.graham@lubs.leeds.ac.uk

Dr. Gary Graham is an associate professor of operations and supply chain management at Leeds University, and a visiting research scholar at MIT Future Freight Lab. He has authored three books, 34 research papers, and his grant income (EU, EPSRC and ESRC) totals £1.13m.

Dr Anita Greenhill, Alliance Manchester Business School, Booth Street West, Manchester M15 6PB, UK; +44 161 2003508; a.greenhill@manchester.ac.uk

Dr. Anita Greenhill is a senior lecturer at the University of Manchester. Her research interests are in the areas of Networked Usage of Technology within Community, Organisational and Business settings. Dr Greenhill's research focus on building a better understanding of crowdsourcing and Internet participation; Community usage/resilience and its use of social media including twitter; technologically enable work; spatiality; and Web Usage in Organisations.

Professor Chris Lintott; University of Oxford Department of Astrophysics, Keble Road, Oxford, OX1 3RH, UK; +44 1865 273638; chris.lintott@physics.ox.ac.uk

Prof. Chris Lintott is a professor of astrophysics at the University of Oxford, where he is also a research fellow of New College. He is the principal investigator for Zooniverse. His research focuses on the development of sophisticated platforms for citizen science pattern recognition and anomaly detection (as well as galaxy evolution).

Dr Karen Masters; University of Portsmouth Institute for Cosmology and Gravitation, Dennis Sciama Building, Burnaby Road, Portsmouth, PO1 3FX, UK; +44 2392 843137; karen.masters@port.ac.uk

Dr. Karen Masters is a Reader in extragalactic astronomy at the Institute of Cosmology and Gravitation, University of Portsmouth. She uses information from large surveys of the sky to search for clues about how galaxies formed and evolve over cosmic time. Dr. Masters is the Project Scientist of the Galaxy Zoo Science Team.

Dr Jamie Woodcock; Department of Management, London School of Economics and Political Science, Houghton Street, London, WC2A 2AE, UK; +44 20 7955 6823; j.woodcock@lse.ac.uk.

Dr Jamie Woodcock is a Fellow in Management the London School of Economics and Political Science. His areas of research interest include the labour process, workplace resistance, management organisation and critical theory.

ACKNOWLEDGEMENTS

This work was partially supported by funding from the EPSRC and New Economic Models in the Digital Economy (NEMODE) network as part of the Volunteer and Crowdsourcing Economics project (EP/K039784/1). The survey work has been reviewed and approved by the University of Portsmouth Business School Ethics Committee (REF: E315).

Doing good online: The changing relationships between motivations, activity and retention among online volunteers

ABSTRACT

Advances in Internet technology are making it possible to volunteer online through participation in research-based activities supporting non-profit and charitable organisations. Using survey data from a representative sample of contributors, this study investigates motivations to volunteer across a sample of five such online projects using the Volunteer Functions Inventory. We explore relationships between these motivations and actual recorded measures of both volunteer activity and retention. We also use quantile regression analysis to investigate the extent to which these motivations change at different stages in the volunteer process. Our results show that activity and retention tend to associate significantly and positively with the understanding and values motivations, as well as significantly and negatively with the social and career motivations. We also find the importance of motivations changes significantly among different percentiles of volunteer engagement. For some motivations, especially understanding, the nature of these changes is markedly different between activity and retention.

Keywords: Online; Volunteering; Motivations; Volunteer Functions Inventory

1: Introduction

While much has been written concerning the growth of the digital economy and its impact upon commercial activities, relatively little attention has been paid to the effects of digitisation upon the voluntary sector. Internet-based volunteering projects are truly many and varied, but typically involve aggregation of input from large numbers of contributors working together towards a common goal. Possibly the best-known among such projects is Wikipedia, a free online encyclopaedia co-created and maintained exclusively by volunteer contributors. The rise of these new forms of online volunteering may have significant implications for the academic study of volunteering, not least because online volunteering may serve to complement more conventional offline forms of the activity (Ihm, 2017). However, an overwhelming majority of extant theory and evidence remains focused exclusively on conventional offline forms of volunteering. As a consequence, there is a need to develop a more detailed understanding of the effects of digitisation on volunteer activity and retention, as well as the ways in which these change at different stages of the online volunteering process.

We address this deficiency in the literature through the analysis of a dataset collected from five different online volunteering projects hosted by the Zooniverse, a web-based portal which allows citizens to participate in collaborative research activities managed by teams based in museums, universities and other non-profit organisations. Our analysis combines results from a large-scale survey undertaken with a representative sample of registered contributors, containing information on socio-demographic, attitudinal and behavioural characteristics, as well as items appearing on the Volunteer Functions Inventory (VFI). We reconcile this against an extensive database of user interactions in order to examine to examine the extent to which VFI motivations explain variations in actual activity and

retention levels, as well as the degree to which their importance changes at different stages of the volunteering process. Thus, our study uses recorded data on observed behaviours, unlike the vast majority of research into volunteering which relies exclusively upon self-reported activity levels or stated intentions.

The aim of this study is to address two specific research questions. In an approach which is consistent with other prior studies of volunteer motivation in different contexts, Research Question 1 asks *'how do VFI motivations relate to variations in recorded patterns of activity and retention among online volunteers'?* In addressing this question, we specify a series of multiple regressions to model volunteer engagement using a range of activity and retention measures as dependent variables. Our independent variables include the set of factor scores relating to items from the VFI alongside other socio-demographic, attitudinal and behavioural controls captured by our survey data.

Following this analysis, our second research question aims to offer a deeper and more nuanced insight into the changing nature of volunteer motivations at different stages of activity and retention, thus contributing to both theoretical understanding and empirical evidence on the subject. Research Question 2 asks '*to what extent do the importance of VFI motivations change at different stages of the volunteering process'?* In addressing this question, we employ a series of quantile regressions which demonstrate whether the strength and nature of the relationships between individual motivations and volunteer engagement change among volunteers who contribute higher levels of effort and/or over longer periods. In other words, while a typical regression analysis might demonstrate a particular motivation to be an important determinant of variations in volunteer engagement, quantile regressions demonstrate whether those motivations are more or less important in explaining variations in behaviour at higher or lower percentiles of both activity and retention.

2: Online Volunteering and the Zooniverse

One of the best known voluntary crowdsourcing platforms in the field of non-commercial research is the Zooniverse, a collection of more than forty active online research projects powered by volunteer contributors (Fortson *et al*., 2012). Projects hosted by the Zooniverse represent an innovative response to challenges posed by increasingly large and complex data sets, especially in cases where it is difficult or impossible for computers to interpret such data automatically. Zooniverse projects ask for input from human volunteers to interpret or 'classify' these datasets with the ultimate objective of helping teams of professional researchers in non-profit organisations and charities address a range of specific research questions. For example, the first and one of the best-known Zooniverse projects, Galaxy Zoo, presents users with a series of images of deep-space galaxies requiring classification according to a set of pre-defined criteria relating to their shape (Lintott *et al*., 2008). The resultant analysis of data gathered from volunteers is helping astrophysicists develop a better understanding of the evolution of galaxies. Other examples of Zooniverse projects include Cell Slider, which asks volunteers to analyse the properties of cancer cells to help Cancer Research UK develop treatments and Wildcam Gorongosa, where volunteers classify images of animals from camera traps stationed around the Gorongosa National Park(1). The Zooniverse has been hugely successful since its launch in 2010 and now has more than 1.3 million registered participants. On average across each individual Zooniverse project, volunteers contribute the amount of information that it would take a professional researcher 34 full-time years working alone to complete (Cox *et al*., 2015).

Our study is based on the analysis of user motivations and behaviours within a number of Zooniverse projects in the areas of astrophysics and ecology; namely, Galaxy Zoo, Planet Hunters, Seafloor Explorer, Snapshot Serengeti and Penguin Watch. Figure 1 contains

screenshots of the different online interfaces for these projects. In each case, a volunteer is either asked to answer a series of questions about the properties of an image they see, or are asked to point and click to areas of an image relating to content of particular research interest. Sophisticated algorithms are subsequently applied to convert the large quantity of volunteer data supplied for each individual image into a consensus solution that can be used for further research. The high quality of the research data generated by Zooniverse projects is highlighted by more than 100 publications[(2)] in peer-reviewed academic journals that have only been possible as a result of input from volunteers, many of whom are thanked in the author acknowledgements or even credited as formal co-authors.

[Figure 1 about here]

Activities hosted or linked by portals such as the Zooniverse or Crowdcrafting can be more broadly termed as online 'citizen science' projects, which differ from other online initiatives in a number of ways. Although based around a model of micro-tasking, Zooniverse projects differ from paid initiatives such as Amazon's Mechanical Turk and other innovation-based online crowdsourcing initiatives such as Innocentive which offer prizes and other material incentives to participants. Online citizen science projects are more akin to voluntary-imbedded online initiatives, such as Open Source Software (OSS) development and content-driven projects such as Wikipedia, but are further distinct in terms of their specific remit to involve contributors in real scientific research. Within the specific field of citizen science, projects hosted on the Zooniverse are best thought of as 'volunteer thinking' initiatives, where users are presented with data and are trained to analyse it according to research protocols (Jennett *et al*., 2016). Volunteer thinking projects differ from other forms of citizen science, such as 'volunteer computing', where users install software on their computers which automatically takes advantage of unused computing power (such as SETI@Home), or

'participatory sensing' which involves data gathering and submission, usually involving mobile phone apps (such as Noisetube).

A small number of prior studies have specifically investigated the motivations of contributors to Zooniverse projects. Among the most cited of these, Raddick *et al.* (2010) investigates the motivations of Galaxy Zoo participants, first holding interviews to determine motivations and subsequently grouping these into discrete categories and surveying a larger sample of users in a follow-on study (Raddick *et al.*, 2013). They found that 'being excited by the opportunity to make an original contribution to science' was the most commonly stated motivation to participate. However, these studies suffer from a combination of focus on a single project, a high likelihood of selection bias in the composition of the survey sample and an absence of any investigation into the relationship between identified motivations and patterns of volunteer activity and retention. In addition to overcoming each of these limitations, we also undertake a quantile regression analysis in order to establish the extent to which motivations to participate in these projects change at different stages of the volunteering process. This complements the more qualitative approach used in studies such as Rotman *et al.* (2012; 2014) to assess the extent to which motivations to participate in citizen science projects change over time.

3: Literature Review & Conceptual Framework

Much of the extant literature relating to the motivation of online volunteers relates to participation in OSS development, such as the Linux kernel. Although earlier work tends to highlight intrinsic motivations such as altruism as the primary driver of participation (Haruvy *et al*., 2003; Lakhani & Wolf, 2005), subsequent studies identify a much wider and more

complex combination of motivations. These include various extrinsic factors (Lerner & Tirole, 2005), particularly social and community needs based around a desire for interaction or reputational enhancement (Jannsen & Huang, 2008; Xu *et al*., 2009). It has also been noted by various authors that career motivations tend to be particularly prevalent among contributors to OSS projects (Hertel *et al*., 2003; Wu *et al*., 2007), perhaps given the high proportion of contributions from IT professionals (Bitzer & Geishecker, 2010). By contrast, contributors to Wikipedia are shown to be motivated by more altruistic factors (Yang & Lai, 2010), although Xu & Li (2015) find only participation in community aspects of Wikipedia to be intrinsically motivated. They argue instead that the contribution of content is largely motivated by extrinsic factors, such as self-development and reciprocity.

A distinctive feature of our study is the use of the functional approach to human behaviour as a theoretical basis for our analysis. This approach is largely based on the theories of Smith *et al.* (1958) and Katz (1960), which assert that volunteers are motivated by a desire to satisfy various combinations of social and psychological goals. The most well-known and complete metric used to measure and interpret volunteer motivation is the Volunteer Functions Inventory (VFI), a formal instrument consisting of six distinct items that was pioneered by Clary *et al.* (1996). These six motivations are; Protective (a means to shield or escape from problems); Enhancement (a means to feel better about oneself); Social (a means to interact with people and expand social networks); Values (a means to express personal values and contribute to important causes); Understanding (a means to gain new perspectives and to learn) and Career (a means to build skills and connections to enhance one's career). These six items have been shown to be robust and consistent when applied across different cohorts of volunteers, as well as across time and different forms of volunteering (Clary *et al*., 1998).

The VFI has subsequently seen widespread use in analysing the motivations of volunteers for a number of activities and organisations around the world. Although results tend to differ depending on the particular context (Stukas *et al.*, 2009), a majority of studies highlight the importance of the 'other-oriented' motivations of Values, Understanding and Social in explaining variations in volunteer activity (e.g. Gage & Thapa, 2012; Stukas *et al.*, 2014), with lower importance attached to the more 'self-oriented' motivations of Protective, Enhancement and Career (Planalp & Trost, 2009; Agostinho & Paco, 2012). By contrast, volunteer retention has been shown to associate significantly and positively with understanding and social motivations (Ferreira *et al*., 2015; Hyde *et al*., 2016), and negatively with the career motivation (Tschirhart *et al*., 2001; Garner & Garner, 2011). The only prior study of which we are aware that has formally applied the VFI in in the context of online volunteering is by Nov (2007), which finds that the more altruistic motivations of values and understanding do a better job of predicting variations in self-reported activity levels among Wikipedia contributors. However, these findings are limited due to reliance on a self-selecting group of survey respondents that are not necessarily representative of the population being studied.

By comparison, there are relatively few studies in the volunteering literature that adopt a quantitative approach to investigate the changing motivations of volunteers at different stages of the volunteering process. Studies investigating these issues tend to be largely qualitative in nature and based on the analysis of relatively small samples (e.g. Kelly *et al*., 2005). A small number of studies investigate partly-related issues, such as variations in motivations over time due to age and personal circumstance (Nesbit, 2012) or differences in motivations between current and former volunteers (Hustinx, 2010). One exception is the study undertaken by Finkelstein (2008), which adopts a longitudinal, quantitative approach to

investigate changing motivations within a sample of volunteers over a 12-month interval. However, this study is limited due to its reliance on simple correlation coefficients between key variables. Analysing this issue using quantile regression analysis allows us to assess the changing nature of motivations to volunteer at different stages of the process *ceteris paribus* through the inclusion of appropriate control variables.

4. Data

4.1. Data Sources

This study combines data from two main sources, the first of which is a survey of volunteers for five different Zooniverse users undertaken during April and May 2015. The survey was entirely web-based, with each individual respondent being e-mailed a unique URL allowing each response to be linked to a particular user account. This approach also allowed us to collect relevant data on both volunteer activity and retention directly from the central Zooniverse database. In terms of volunteer activity, we measure the aggregate amount of analysis supplied by each user (number of 'classifications') for the entire portfolio ofZooniverse projects, as well as the number of individual projects towards which the volunteer has contributed. We further measure the retention of volunteers through the number of unique days where classifications were recorded and the length of time for which each respondent has 'actively' contributed towards projects (the time difference between first and last recorded classifications). A summary of the correlations between these measures of engagement and retention appears as Table 1. As expected, our four measures of engagement and retention are positively correlated, although the degree of correlation between them is mostly weak. The one exception to this is the relatively strong positive correlation between

the number of recorded classifications (activity) and number of unique days on which classifications were recorded (retention). However, we are broadly satisfied that each of these measures represents a somewhat different facet of volunteer behaviour and are sufficiently distinct from one another to warrant separate investigation of each measure.

[Table 1 about here]

After excluding a very small number of obvious outliers, our final dataset comprises a total of 1,915 survey responses. The comparison shown in Figure 2 indicates that is that the distribution of activity (number of classifications submitted) among our sample of survey respondents broadly matches the distribution observed for the whole population of Zooniverse volunteers quite closely, with only a slight discrepancy in terms of the number of survey respondents supplying a single classification compared with the equivalent proportion among the entire population of Zooniverse volunteers.

[Figure 2 about here]

4.2. Descriptive Statistics

Descriptive statistics for the sample of survey respondents are presented in Table 2. As can be seen from our various measures of engagement, the distribution of voluntary contributions made by Zooniverse users is highly skewed. A raw count of the total number of Zooniverse classifications recorded among the sample shows a mean of around 2,733 per user, versus a median of 260 and a standard deviation around 7 times larger than the mean. A further investigation of the distribution of classification activity among users show that the top 5% of volunteers by overall classification-count provide around 70% of the total recorded classifications; in other words, providing more than twice the number of classifications compared with the other 95% of users combined. The same skewed distribution of activity is

also observed for the other measures of volunteer activity and retention summarized in Table 2. Overall, we can see that a vast majority of volunteers supply a relatively small number of classifications over a very short period of time; usually a handful of sessions lasting only a few hours in total.

[Table 2 about here]

These descriptive statistics further show that our dataset contains a roughly equal proportion of male and female respondents. The average educational attainment among the sample is relatively high, with around 67% of survey respondents holding a Bachelor's degree or higher. Of this proportion, around 36% hold a Master's Degree and around 12% have doctorate-level qualifications. Around half of the total number of respondents hold their highest qualifications in science-related subjects. Additionally, our sample appears to consist of a relatively high proportion of white respondents (87%) living in cities (66%). Respondents to our survey also appear to be reasonably affluent; just over half own their own homes, with an average annual income of just over $40,000 per annum.

4.3. Volunteer Functions Inventory Motivations

In addition to socio-demographic information, the survey also collected information on motivations to volunteer. We employ a subset of three out of the five questions under each heading of the VFI selected on the basis of the strength of correlation with the relevant underlying factor score reported by Clary *et al.* (1998). In most cases, the wording for each question needed to be modified only slightly to make it specific to the particular context of online volunteering via the Zooniverse. Survey respondents were asked to indicate the extent to which they agreed or disagreed with each of these statements on a 7-point Likert scale and

were each presented with the statements in a random order. A confirmatory factor analysis was undertaken to verify the goodness of fit of our Likert scale data to the standard six-factor VFI solution. As per the guidelines suggested by Schumacker & Lomax (2004), the confirmatory factor analysis allowing for correlation between latent factors generally indicates good levels of construct validity, with a low RMSEA ($0.04 < 0.05$) indicating a good absolute fit, a high CFI ($0.96 > 0.90$) indicating a good incremental fit and a ratio of Chi Squared / Degrees of Freedom ($4.73 < 5$) indicating a parsimonious fit to the data.

A summary of the variables used in this process is presented in Table 2, as well as the loadings for each individual attitudinal response against the respective factor score. Cronbach's Alpha values for each factor are shown to be appropriately high, except in the case of the Values motivation. This lower level of internal consistency may be a consequence of our decision to replace several of the standard values questions we felt were inappropriate in this context, such as being 'concerned over others less fortunate than oneself', with alternatives asking about the extent to which the respondent believes that scientific research benefits society and whether scientific research receives adequate funding. A number of prior studies (e.g. Liao-Troth, 2005, Marta *et al*., 2006; Francis, 2011) have previously found that factors based upon the values motivation have associated with the lowest levels of internal consistency among VFI motivations, often with alpha values close to or below conventional thresholds. An exploratory factor analysis performed on our dataset identifies responses to the three values questions as satisfactorily loading on the same factor using conventional measures (e.g. Eigenvalue > 1, accounting for 6% of the 65% cumulative total variance explained by all six VFI factors). However, the results presented in subsequent sections relating to this factor score should nonetheless be considered exploratory and treated with a degree of caution.

[Table 2 about here]

5. Analysis

5.1. Research Question 1: How do VFI motivations relate to variations in activity and retention among online volunteers?

The previous section outlined the process undertaken to generate factor scores for each item on the VFI. These scores are used as explanatory variables in a regression analysis using as dependent variables observed measures of volunteer activity in model specifications (i) – (ii) and volunteer retention in model specifications (ii) – (iv). The results of these regressions are presented in Table 4. Other variables from our survey are also included in these specifications in order to control for variations in individual socio-demographics and lifestyle choices, although it should be noted we do not report all coefficient estimates for these control variables in order to conserve space.

[Table 4 about here]

Although we use four different measures of volunteer engagement as dependent variables, the broad conclusions are similar across model specifications, indicating that our results are robust to different measures of both volunteer activity and retention. The most significant positive associations between engagement levels and motivation appears to relate to understanding and values, where our coefficient estimates are found to be relatively large and statistically significant in almost every specification. These regression results therefore show clear evidence that the most engaged participants are primarily motivated by a desire to enhance their levels of knowledge and understanding, as well as adherence to (scientific) values.

We also show evidence of a generally positive association between volunteer activity levels and the protective motivation, although this relationship does not appear to hold for volunteer retention. This means that volunteers who contribute more actively, but not necessarily more frequently, tend to be at least partly motivated by a desire for escapism. Our results also demonstrate strong negative relationships between both the career and social motivations and all of our measures of volunteer activity and retention, suggesting that the most active and longer-serving participants are significantly less motivated by the possibility to enhance their careers or to socialise with other volunteers compared with others. We find no significant association between the enhancement motivation and any of our measures of volunteer activity or retention.

Surprisingly, we show only very limited association between our socio-demographic controls and either the engagement of online volunteers; particularly with respect to age, gender, ethnicity, income and education. The only constraint on engagement seems to be the relationship status of the participant, with weakly significant reductions in engagement observed among respondents in a committed relationship compared with those who are single. As expected, the time control variable (Duration) also indicates that volunteers who have held accounts for longer periods tend to have been more active over time in terms of classification activity and number of visits. Altogether, taken across all of our model specifications, the six factor scores generated from the VFI motivation items seem to do a much better job of explaining variations in volunteer activity and retention than our set of socio-demographic controls. This indicates that individual level motivations are more powerful predictors of variations in the activity and retention of online volunteers than those reflecting respondent characteristics and lifestyle choices.

5.2. Research Question 2: What role do VFI motivations play at different stages of the volunteer process?

In order to investigate the extent to which these identified motivations might play different roles at different stages in the volunteer process, we undertake a series of quantile regressions, the results of which can be found in Table 5a (activity) and Table 5b (retention). Whereas standard linear regression techniques estimate changes in the conditional mean of the dependent variable given certain values of one or more explanatory variables, quantile regression estimates changes in one or more specified percentiles of the dependent variable. For example, when the quantile is set to 0.50 (the median), the effect of each explanatory variable is estimated in relation to changes in the conditional median level of volunteer activity/retention. By comparing coefficient estimates at different quantiles (in this case, 0.25, 0.50 and 0.75), it is therefore possible to determine whether some percentiles of volunteers are affected to a greater or lesser extent by different motivations. In other words, while the standard linear regressions outlined in the previous section indicate the motivations that are important determinants of volunteer activity and retention, quantile regressions can demonstrate whether these motivations are more or less important for volunteers with higher or lower levels of activity and retention. The results demonstrate pronounced changes in the relationships between several motivations and volunteer activity/retention at different stages of the volunteer process.

[Tables 5a and 5b about here]

With respect to our measures of volunteer activity, the results in Table 5a clearly show that the understanding motivation associates even more strongly and positively with volunteering at higher percentiles of activity. We also observe a similar, albeit slightly weaker, increasing

importance the protective motivation, while that the career motivation associates even more negatively with activity at later stages of the volunteer process; the effect of the latter appearing to be stronger with respect to number classifications (specification i) than number of projects (specification ii). Conversely, we see a reduction in the negative association between the social motivation at higher percentiles of activity. Taken together, it appears that the understanding and protective motivations become more important in explaining increases in activity among more active volunteers, while career motivations appear more strongly negatively associated with activity levels at later stages in the volunteer process. By contrast, social motivations appear to associate less negatively with higher percentiles of activity.

With respect to our measures of volunteer retention, the results in Table 5b demonstrates that the two motivations that associate most significantly with volunteer engagement overall, understanding and values, become less important at higher percentiles of volunteer retention. In the case of the former, a comparison with the results presented in Table 5a suggests that while understanding becomes a more important motivation at higher percentiles of volunteer activity, it seems to become less of an important motivation at higher percentiles of volunteer retention. Similarly, the importance of the values motivations appears to diminish at higher percentiles of one of our measures of retention (active period), but not the other (unique days). Altogether, this implies a marked contrast in the development of motivations between volunteers who contribute more intensively (activity) versus those who contribute over a sustained period (retention). Additionally, we show mixed evidence on the changing nature of the career motivation, which becomes more unimportant at higher percentiles of one measure of retention (unique days), but less unimportant according to the other (active period). Given that the other coefficient estimates are shown to be statistically equivalent to zero for all measured percentiles, we find no evidence to suggest that any other motivations

have differing effects upon volunteer retention at different stages of the volunteering process. It therefore appears that retention levels are driven less by understanding at higher percentiles of volunteer retention, with slightly weaker evidence pointing to a reduction in the importance of values and career motivations.

6. Discussion

Overall, the results presented above suggest that activity and retention levels for contributors to our sample of online volunteering projects tend to associate positively and significantly with the understanding and values motivations, as well as significantly and negatively with the social and career motivations. The protective motivation appears to associate somewhat positively with volunteering activity (though not retention), while the social motivation is also found to become less unimportant at higher percentiles of both volunteer activity and retention. Further, we also show the effects of understanding and protective motivations to be more important among more active percentiles of volunteers, while understanding and, to some extent, values appear to diminish in importance among longer-serving percentiles of volunteers. These findings suggest that the importance of motivations can change significantly at different stages of the volunteering process, while the nature of these changing motivations can be markedly different between activity and retention. Broadly speaking, our findings suggest that online volunteering in the context of citizen science might initially be driven more by the understanding and values motivations, with these eventually diminishing in importance and the protective motivation subsequently increasing in importance according to at least some of our measures of volunteer engagement.

Although the importance of various motivations is shown to change at different stages of the volunteer process, our results consistently show that learning is an important component of the process of volunteering in this particular context, demonstrating strong positive associations with almost all measures of activity and retention and becoming more important among higher percentiles of the sample by activity. However, the lack of positive association between volunteer engagement and educational attainment suggests that, even though the sample of online volunteers seems to be relatively highly educated compared with the rest of the population, those with the highest existing levels of education do not appear to contribute the most voluntary effort and information towards these projects. In a practical sense, online volunteering and citizen science projects should therefore offer clear opportunities for learning to incentivise participation, as well as encouraging and promote opportunities for learning among all participants, particularly among the most active contributors. Given recent evidence that learning actually does occur among more active participants in citizen science projects (Masters *et al*., 2016), this further signals how opportunities for learning could represent an effective means by which to not only motivate participation, but also to increase scientific literacy and knowledge within society.

Although we do not find evidence of a strong, positive association with the protective motivation across the whole of our sample, the quantile regression results show that this motivation becomes more important among more active participants. For online citizen science initiatives such as the Zooniverse, this may be an important consideration from a project design and management perspective, given that users at later stages of the volunteering process appear to be motivated to a greater extent by a desire for escapism, whereas users at an earlier stage are not. This may suggest that the provision of more complex and involved tasks might be appropriate to present to highly active participants,

while simpler tasks requiring casual engagement would be more appropriate for newer and less active volunteers.

Additionally, the consistent evidence of a negative association between the social motivation and our measures of activity and retention implies some degree of substitutability between social interaction and cognitive input into these online projects. Although this may partly be a result of the way in which these projects have been designed (volunteer classifications need to be independent of one another to ensure statistical validity of the findings), the Zooniverse does offer extensive facilities for interaction and discussion within its community of volunteers via the 'Talk' feature. Our results imply that participants typically want to either socialise or classify, although our quantile regressions indicate that socialisation may become less of a direct substitute to both activity and retention at later stages in the volunteering process. The Zooniverse may therefore benefit from more specifically targeting their social and community elements to participants who are more engaged in terms of activity and retention. Additionally, career motivations are shown to associate negatively with both activity and retention, which broadly indicates that those with careers in related fields are less likely to volunteer in this particular context. This highlights how more active online citizen science volunteers are predominantly made up of non-professionals lacking career aspirations in scientific fields, which stands in stark contrast to the findings from studies relating to OSS development cited earlier.

By way of placing our findings in context with others relating to online volunteering, the importance of understanding and values as motivations for online volunteer activity is consistent with many other findings appearing elsewhere in the literature on OSS, such as Ye & Kishida (2003) and Fang & Neufeld (2009). However, although other studies have also highlighted the changing nature of motivation among contributors to other online projects,

not all authors agree on which and how motivations change between early and longer-standing contributors. For example, while David & Shapiro (2008) find that the flow of individuals from smaller to larger projects over time is primarily driven by a desire to learn and enhance skills, other studies such as Shah (2006) argue that more 'serious' contributors are less motivated by values and are driven to a greater extent by social and community-based issues such as reciprocity, feedback and reputation. Additionally, Budhathoki & Haythornthwaite (2013) suggest that learning and social motivations are more prevalent among more 'serious' contributors to the OpenStreetMap project, while Curtis (2015) finds the most important initial motivations among contributors to the Foldit citizen science project relate to values, while heavier contributors express a greater desire for social interaction and intellectual challenge. Our findings suggesting that activity and retention levels among online volunteers do not seem to vary according to sociodemographic characteristics also stands in contrast with the findings of Abdelkader (2017), who finds that factors such as age, gender and educational attainment significantly affect participation in online volunteering activities connected with education.

Overall, although the precise findings of these studies are dependent on the particular context, they do generally seem to point towards a trend of initial motivation by values (or similar), which changes over the course of a participant's association more towards being motivated by factors associated with social and understanding. While our results also broadly reinforce this pattern, ours differ in observing an increase in the importance of the protective motivation among higher percentiles of activity and a decreasing importance of understanding among higher percentiles of retention. Our study also differs in our finding that the career motivation associates more negatively with activity and less negatively with retention among higher percentiles of our sample of volunteers. Given these differences, we

suggest that more research is needed that makes use of the VFI framework alongside techniques such as quantile regression analysis to establish whether these particular findings also hold in other online volunteering and citizen science contexts.

Although our dataset consists of observations relating to contributors to five distinct online volunteering projects, the principal limitation of this study if that all of the sampled projects are hosted on the Zooniverse portal. The extent to which these findings are therefore generally applicable to other forms of voluntary-imbedded online initiatives such as OSS development or Wikipedia, is therefore debatable. Nonetheless, our study does offer a number of valuable insights into the motivations for engagement in this novel form of online volunteering, as well as the way in which these motivations change at different stages of the volunteering experience. This is likely to be particularly important given the rapid growth of this type of online volunteering project and the likelihood of increased competition to recruit and retain volunteers in the future (Kargh, 2016). Our findings are therefore not only of value to existing citizen science activities, but also other research-related volunteering initiatives, especially the increasing number of charities and non-profit organisations seeking to engage volunteers in online settings (Saxton & Guo, 2011).

7. Conclusions

This study has introduced an underexplored form of online volunteering known as citizen science, which provides opportunities for citizens to engage in research-related activity and analysis for a wide variety of non-profit and charitable organisations. The profile and motivations of these volunteers is explored through access to an extensive database from the portfolio of online projects hosted by the Zooniverse portal, containing actual, observed

measures of both voluntary activity and retention. This contrasts with the majority of studies exploring variations in volunteer activity and retention which rely on self-reported activity levels and intentions to continue volunteering in the future. We supplement this information with the results from an online survey undertaken with a representative group of volunteers across five different online projects. This crucially allows us to measure and control for a variety of socio-demographic characteristics, as well as motivations to volunteer expressed using the VFI framework.

Following our analysis of these data, we find only limited evidence to suggest that socio-demographic characteristics explain much of the variation in the levels of engagement observed between online volunteers; either in terms of activity levels or retention. By contrast, the motivations to volunteer encapsulated by the VFI appear to do a much better job of explaining variations in engagement. More specifically, we show that the understanding and values motivations generally have the strongest positive association with volunteer engagement, while the career and social motivations are shown to associate the most negatively. Given the strong, positive association between the understanding and values motivation and all measures of activity and retention, we conclude that opportunities for learning relating to the core values of the volunteering initiative appear to represent the most effective incentive to volunteer in this particular context. It therefore appears that online volunteering for Zooniverse and other online citizen science projects may be more concerned with knowledge creation and human capital enhancement than more traditional forms of offline volunteering.

Additionally, our study demonstrates how these motivations change at different stages of the online volunteering process. Broadly, we show how the importance of the understanding motivation tends to increase in importance as online volunteers become more active, but

diminishes in importance for longer-serving volunteers. By comparison, the importance of the values motivation appears to diminish as some measures of volunteer engagement increase, while the relative importance of the protective and social motivations appears to increase. We therefore suggest that volunteering for online citizen science projects might initially be motivated to a greater extent by the understanding and values motivations, with the protective and social motivations subsequently increasing in relative importance at later stages of the volunteering process. On the basis of these findings, we make a number of recommendations regarding the design of online citizen science projects, including offering clear opportunities for learning, varying the difficulty of task according to experience and targeting the social and community elements of projects to the most engaged volunteers.

Taken together, these results afford a detailed insight into the changing nature of motivations of contributors to these novel forms of online volunteering, as well as the contrast between evolving motivations between more active and longer-serving volunteers. These changes and contrasts should be further researched through the more widespread use of more sophisticated empirical techniques, such as quantile regression, in order to better understand the changing nature of these motivations among other groups of volunteers in both online and offline settings.

Endnotes

(1) A full list of current Zooniverse projects can be found at http://www.zooniverse.org.

(2) A full list of all peer-reviewed publications resulting from Zooniverse projects can be found at https://www.zooniverse.org/publications.

References

Abdelkader, O. A. (2017). Significant concerns influence online pro bono volunteering of faculty members. *Computers in Human Behavior*, 73, 547-553.

Agostinho, D., & Paço, A. (2012). Analysis of the motivations, generativity and demographics of the food bank volunteer. *International Journal of Nonprofit and Voluntary Sector Marketing*, 17(3), 249-261.

Bitzer, J., & Geishecker, I. (2010). Who contributes voluntarily to OSS? An investigation among German IT employees. *Research Policy*, 39(1), 165-172.

Budhathoki, N. R., & Haythornthwaite, C. (2013). Motivation for open collaboration crowd and community models and the case of OpenStreetMap. *American Behavioral Scientist*, 57(5), 548-575.

Clary, E. G., Snyder, M. & Stukas, A. A. (1996). Volunteers' motivations: Findings from a national survey. *Nonprofit and Voluntary Sector Quarterly*, 25(4), 485-505.

Clary, E. G., Snyder, M., Ridge, R. D., Copeland, J., Stukas, A. A., Haugen, J., & Miene, P. (1998). Understanding and assessing the motivations of volunteers: a functional approach. *Journal of Personality and Social Psychology*, 74(6), 1516.

Cox, J., Oh, E. Y., Simmons, B., Lintott, C., Masters, K., Graham, G., Greenhill, A. & Holmes, K. (2015). How is success defined and measured in online citizen science? A case study of Zooniverse projects. *Computing in Science and Engineering*, 17(4), 28-41.

Curtis, V. (2015). Motivation to Participate in an Online Citizen Science Game: A Study of Foldit. *Science Communication*, 37(6), 723-746.

David, P. A., & Shapiro, J. S. (2008). Community-based production of open-source software: What do we know about the developers who participate? *Information Economics and Policy*, 20(4), 364-398.

Fang, Y., & Neufeld, D. (2009). Understanding sustained participation in open source software projects. *Journal of Management Information Systems*, 25(4), 9-50.

Ferreira, M.R., Proença, T., & Proença, J.F. (2015). Volunteering for a lifetime? Volunteers' intention to stay in Portuguese Hospitals. *VOLUNTAS*, 26(3), 890-912.

Finkelstein, M.A. (2008). Predictors of volunteer time: The changing contributions of motive fulfillment and role identity. *Social Behavior and Personality*, 36(10), 1353-1363.

Fortson, L., Masters, K., Nichol, R., Edmondson, E. M., Lintott, C., Raddick, J., & Wallin, J. (2012). Galaxy Zoo. *Advances in Machine Learning and Data Mining for Astronomy*, 213-223.

Francis, J.E. (2011). The functions and norms that drive university student volunteering. *International Journal of Nonprofit and Voluntary Sector Marketing*, 16(1), 1-12.

Gage, R.L. & Thapa, B. (2012). Volunteer motivations and constraints among college students analysis of the volunteer function inventory and leisure constraints models. *Nonprofit and Voluntary Sector Quarterly*, 41(3), 405-430.

Garner, J.T., & Garner, L.T. (2010). Volunteering an Opinion Organizational Voice and Volunteer Retention in Nonprofit Organizations. *Nonprofit and Voluntary Sector Quarterly*, *40*(5), 813-828.

Haruvy, E., Prasad, A., & Sethi, S. P. (2003). Harvesting altruism in open-source software development. *Journal of Optimization Theory and Applications*, 118(2), 381-416.

Hertel, G., Niedner, S., & Herrmann, S. (2003). Motivation of software developers in open source projects: An Internet-based survey of contributors to the Linux kernel. *Research Policy*, 32(7), 1159-1177.

Hustinx, L. (2010). I quit, therefore I am? Volunteer turnover and the politics of self-actualization. *Nonprofit and Voluntary Sector Quarterly*, 39(2), 236-255.

Hyde, M. K., Dunn, J., Bax, C., & Chambers, S. K. (2016). Episodic volunteering and retention: An integrated theoretical approach. *Nonprofit and Voluntary Sector Quarterly*, 45(1), 45-63.

Ihm, J. (2017). Classifying and Relating Different Types of Online and Offline Volunteering. *VOLUNTAS*, 28(1), 400-419.

Janssen, O., & Huang, X. (2008). Us and me: Team identification and individual differentiation as complementary drivers of team members' citizenship and creative behaviors. *Journal of Management*, 34(1), 69-88.

Jennett, C., Kloetzer, L., Schneider, D., Iacovides, I., Cox, A., Gold, M., Fuchs, B., Eveleigh, A., Mathieu, K., Ajani, Z. & Talsi, Y. (2016). Motivations, learning and creativity in online citizen science. *Journal of Science Communication*, 15(3), 1-23.

Katz, D. (1960). The functional approach to the study of attitudes. *Public Opinion Quarterly*, 24(2), 163-204.

Kelley, M.S., Lune, H., & Murphy, S. (2005). Doing syringe exchange: Organizational transformation and volunteer commitment. *Nonprofit and Voluntary Sector Quarterly*, 34(3), 362-386.

Kragh, G. (2016). The motivations of volunteers in citizen science. *Environmental Scientist*, 25(2), 6.

Lakhani, K., & Wolf, R.G. (2005). Why hackers do what they do: understanding motivation and effort in free/open source software projects. In: Feller, J., Fitzgerald, B., Hissam, S. & Lakhani, K. (Eds.), *Perspectives on Free and Open Source Software*. MIT Press.

Lerner, J., & Tirole, J. (2005). The economics of technology sharing: Open source and beyond. *The Journal of Economic Perspectives*, 19(2), 99-120.

Liao-Troth, M. A. (2005). Are they here for the long haul? The effects of functional motives and personality factors on the psychological contracts of volunteers. *Nonprofit and Voluntary Sector Quarterly*, 34(4), 510-530.

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. & 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), 1179-1189.

Marta, E., Guglielmetti, C., & Pozzi, M. (2006). Volunteerism during young adulthood: An Italian investigation into motivational patterns. *VOLUNTAS*, 17(3), 221.

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 1-33.

Nesbit, R. (2012). The influence of major life cycle events on volunteering. *Nonprofit and Voluntary Sector Quarterly*, 41(6), 1153-1174.

Nov, O. (2007). What motivates wikipedians? *Communications of the ACM*, 50(11), 60-64.

Planalp, S., & Trost, M. (2009). Motivations of hospice volunteers. *American Journal of Hospice and Palliative Medicine*, 26(3), 188-192.

Raddick, M. J., Bracey, G., Gay, P. L., Lintott, C. J., Murray, P., Schawinski, K., & Vandenberg, J. (2010). Galaxy Zoo: Exploring the Motivations of Citizen Science Volunteers. *Astronomy Education Review*, *9*(1), 010103.

Raddick, M. J., Bracey, G., Gay, P. L., Lintott, C. J., Cardamone, C., Murray, P., Schawinski, K., Szalay, A.S. & Vandenberg, J. (2013). Galaxy Zoo: Motivations of citizen scientists .*arXiv preprint arXiv:1303.6886*.

Rotman, D., Preece, J., Hammock, J., Procita, K., Hansen, D., Parr, C., Lewis, D. & Jacobs, D. (2012, February). Dynamic changes in motivation in collaborative citizen-science projects. In Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work (pp. 217-226). ACM.

Rotman, D., Hammock, J., Preece, J., Hansen, D., Boston, C., Bowser, A., & He, Y. (2014). Motivations affecting initial and long-term participation in citizen science projects in three countries. iConference 2014 Proceedings.

Saxton, G.D., & Guo, C. (2011). Accountability online: Understanding the Web-based accountability practices of nonprofit organizations. *Nonprofit and Voluntary Sector Quarterly*, 40(2), 270-295.

Schumacker, R.E., & Lomax, R.G. (2004). *A beginner's guide to structural equation modelling*: *Second edition*. Mahwah, NJ: Lawrence Erlbaum Associates.

Shah, S. K. (2006). Motivation, governance, and the viability of hybrid forms in open source software development. *Management Science*, 52(7), 1000-1014.

Smith, M. B. (1958). Opinions, Personality, and Political Behavior. *American Political Science Review*, 52(1), 1-17.

Stukas, A. A., Worth, K. A., Clary, E. G., & Snyder, M. (2009). The matching of motivations to affordances in the volunteer environment an index for assessing the impact of

multiple matches on volunteer outcomes. *Nonprofit and Voluntary Sector Quarterly*, 38(1), 5-28.

Stukas, A. A., Hoye, R., Nicholson, M., Brown, K. M. & Aisbett, L. (2014). Motivations to Volunteer and Their Associations with Volunteers' Well-Being. *Nonprofit and Voluntary Sector Quarterly*, December, 1-21.

Tschirhart, M., Mesch, D. J., Perry, J. L., Miller, T. K., & Lee, G. (2001). Stipended volunteers: Their goals, experiences, satisfaction, and likelihood of future service. *Nonprofit and Voluntary Sector Quarterly*, 30(3), 422-443.

Wu, C. G., Gerlach, J. H., & Young, C. E. (2007). An empirical analysis of open source software developers' motivations and continuance intentions. *Information & Management*, 44(3), 253-262.

Xu, B., Jones, D. R., & Shao, B. (2009). Volunteers' involvement in online community based software development. *Information & Management*, 46(3), 151-158.

Xu, B., & Li, D. (2015). An empirical study of the motivations for content contribution and community participation in Wikipedia. *Information & Management*, 52(3), 275-286.

Yang, H. L., & Lai, C. Y. (2010). Motivations of Wikipedia content contributors. *Computers in Human Behavior*, 26(6), 1377-1383.

Ye, Y., & Kishida, K. (2003). Toward an understanding of the motivation of open source software developers. In *25th International Conference on Software Engineering Proceedings*, 419-429, IEEE.

Tables and Figures

Table 1: Correlation Coefficients (Dependent Variables)

	Classifications	Number of Projects	Unique Days	Active Period
Classifications	1			
Number of Projects	0.266	1		
Unique Days	0.814	0.422	1	
Active Period	0.115	0.373	0.268	1

Table 2: Descriptive Statistics

Variable	Description	Mean	Std Dev	Median	Min	Max
Measures of Volunteer Activity						
Classifications	Total number of classifications completed by respondent across all Zooniverse projects.	2732.73	19876.97	260	1	580,000
Number of Projects	Number of unique Zooniverse projects for which the respondent has recorded at least one classification.	5.81	5.50	4	1	35
Measures of Volunteer Retention						
Unique Days	Number of unique days on which the respondent supplied classifications.	29.29	89.00	8	1	2,031
Active Period	Difference (measured in days) between the date of the first and last classifications recorded by the respondent.	841.62	809.71	608	1	2,937
Home Project Controls[†]						
Galaxy Zoo (Base)	Respondent contributes to and answered questions relating to the Galaxy Zoo project.	0.299	-	-	0	1
Planet Hunters	Respondent contributes to and answered questions relating to the Planet Hunters project.	0.247	-	-	0	1
Penguin Watch	Respondent contributes to and answered questions relating to the Penguin Watch project.	0.207	-	-	0	1
Seafloor Explorer	Respondent contributes to and answered questions relating to the Seafloor Explorer project.	0.161	-	-	0	1
Snapshot Serengeti	Respondent contributes to and answered questions relating to the Snapshot Serengeti project.	0.086	-	-	0	1
Other Controls						
Duration	Period of time (in days) between the date of first classification and the date of the survey.	1225.43	788.25	1048	145	2,942
Gender (Female)	Dummy variable if respondent indicated their gender to be female.	0.442	-	-	0	1
Age	Respondent's self-reported age in years.	43.843	15.941	44	18	85
Ethnicity (Non-White)	Dummy variable if respondent indicated their ethnicity to be non-white.	0.129	-	-	0	1
Community Type (Rural)	Dummy variable if respondent indicates they live in a rural area.	0.339	-	-	0	1
Income	Respondent's self-reported income in 2015 USD	41,205	62,541	28,220	0	1,200,000
Religious	Dummy variable if respondent indicated belonging to a religious faith	0.298	-	0	0	1
Charity Donations	Sum of respondent's annual charitable donations in 2015 USD	862.732	2919.201	116.800	0	50,000
Paid Work	Number of hours of paid work undertaken by the respondent in a typical week	23.905	20.569	30	0	95
Relationship (Married/Relationship)	Dummy variable if respondent indicates that they are married or involved in a relationship.	0.49	-	-	0	1
Number of children (aged under 12)	Respondent's number of children aged under 12 years.	0.240	0.626	0	0	6
Number of children (aged under 18)	Respondent's number of children aged under 18 years.	0.126	0.424	0	0	4
Number of children (aged 18+)	Respondent's number of children aged over 18 years.	0.602	1.090	0	0	8
Education Level	Highest educational attainment achieved by the respondent (ISCED Category).	6.587	1.689	7	1	9
Parental Education	Highest educational attainment achieved by either of the respondent's parents (ISCED Category).	5.700	2.098	6	1	9
Science Qualifications	Dummy variable reflecting whether the respondent indicated that the highest qualification achieved was in a scientific field.	0.500	-	-	0	1

[†] 'Home project' refers to the individual Zooniverse project for which each respondent records the largest number of classifications.

Table 3: Confirmatory Factor Analysis

Variable	Standardized Regression Weights	Mean	Std Dev	Median	Min	Max
Career (Chronbach's Alpha = 0.817)						
Participating in Zooniverse projects helps me make new contacts that might help my business or career.	0.768	2.355	1.325	2	1	7
Participating in Zooniverse projects allows me to explore different career options.	0.807	2.881	1.573	2	1	7
Participating in Zooniverse projects will help me to succeed in my chosen profession.	0.754	2.531	1.477	2	1	7
Enhancement (Chronbach's Alpha = 0.803)						
Participating in Zooniverse projects increases my self-esteem.	0.820	4.068	1.505	4	1	7
Participating in Zooniverse projects makes me feel better about myself.	0.757	4.587	1.438	5	1	7
Participating in Zooniverse projects makes me feel needed.	0.711	4.343	1.573	5	1	7
Protective (Chronbach's Alpha = 0.727)						
Participating in Zooniverse projects offers a good way to escape from my troubles	0.695	3.576	1.717	4	1	7
Participating in Zooniverse projects makes me feel less lonely.	0.752	2.929	1.510	3	1	7
Participating in Zooniverse projects makes me feel less guilty about doing enough to support worthwhile causes.	0.632	3.424	1.651	4	1	7
Social (Chronbach's Alpha = 0.851)						
Others with whom I am close place a high value on Zooniverse projects.	0.762	2.928	1.496	3	1	7
My friends contribute to Zooniverse projects.	0.810	2.774	1.513	2	1	7
People I know share an interest in Zooniverse projects.	0.865	3.203	1.641	3	1	7
Understanding (Chronbach's Alpha = 0.806)						
Participating in Zooniverse projects lets me learn through direct, hands-on experience of scientific research.	0.731	5.508	1.255	6	1	7
I feel the Zooniverse allows me to gain a new perspective on scientific research.	0.757	5.496	1.204	6	1	7
Zooniverse projects help me learn about science.	0.813	5.625	1.116	6	1	7
Values (Chronbach's Alpha = 0.402)						
Participating in Zooniverse projects allows me to support a cause I consider to be important.	0.689	6.045	0.994	6	1	7
Scientific research is adequately funded through government taxation.	0.207	3.825	2.129	5	1	6
All of society benefits from scientific research.	0.357	6.378	0.979	7	1	7

Table 4: Multiple Regression Analysis

	VOLUNTEER ACTIVITY		VOLUNTEER RETENTION	
	(i)	(ii)	(iii)	(iv)
	Ln(Classifications)	Ln(Number of Projects)	Ln(Unique Days)	Ln(Active Period)
Career (Factor Score)	-0.393 ***	-0.105 ***	-0.209 ***	-0.242 ***
	(0.091)	(0.040)	(0.057)	(0.094)
Enhancement (Factor Score)	-0.068	-0.039	0.038	-0.064
	(0.122)	(0.047)	(0.075)	(0.123)
Protective (Factor Score)	0.262 *	0.136 **	0.075	0.185
	(0.142)	(0.058)	(0.088)	(0.144)
Social (Factor Score)	-0.107 *	-0.070 ***	-0.070 **	-0.099 *
	(0.056)	(0.023)	(0.035)	(0.057)
Understanding (Factor Score)	0.207 ***	0.088 ***	0.119 **	0.010
	(0.079)	(0.034)	(0.049)	(0.082)
Values (Factor Score)	0.254 **	0.147 ***	0.166 **	0.451 ***
	(0.124)	(0.052)	(0.076)	(0.133)
Gender (Male)	0.054	0.041	-0.006	-0.140
	(0.100)	(0.041)	(0.063)	(0.101)
Ln (Age)	-0.003	-0.073	0.011	-0.102
	(0.171)	(0.069)	(0.107)	(0.172)
Education (Self)	-0.022	-0.001	-0.018	-0.018
	(0.033)	(0.014)	(0.021)	(0.034)
Ethnicity (Non-White)	-0.160	-0.062	-0.129	-0.331 **
	(0.146)	(0.060)	(0.091)	(0.156)
Relationship (Married/Relationship)	-0.264 **	-0.082 *	-0.124 *	-0.164
	(0.114)	(0.047)	(0.072)	(0.117)
Ln (Income)	0.010	0.004	0.001	0.012
	(0.013)	(0.005)	(0.008)	(0.013)
Ln (Duration)	1.066 ***	0.383 ***	0.888 ***	2.168 ***
	(0.073)	(0.032)	(0.047)	(0.065)
Constant Term	-1.568 **	-1.049 ***	-3.861 ***	-9.055 ***
	(0.722)	(0.303)	(0.454)	(0.728)
F Value	18.500 ***	19.040 ***	24.870 ***	58.470 ***
R-Squared	0.189	0.195	0.239	0.358

Significance: * = 90% level, ** = 95% level, *** = 99% level. Other controls included but not reported are as follows: Home Project Controls (Planet Hunters, Penguin Watch, Seafloor Explorer, Snapshot Serengeti, Residence (City), Education (Parents), Science Qualifications, Ln(Paid Work), Ln(Charity Donations), Religion, Number of Children (<12, <18, 18+).

Table 5a: Quantile Regression Analysis (Activity)

Engagement Measure	**(i) Ln(Classifications)**						**(ii) Ln(Number of Projects)**					
Quantile	**0.25**		**0.50**		**0.75**		**0.25**		**0.50**		**0.75**	
Career (Factor Score)	-0.278	**	-0.493	***	-0.443	***	-0.078	**	-0.058		-0.105	**
	(0.140)		(0.117)		(0.144)		(0.036)		(0.054)		(0.053)	
Enhancement (Factor Score)	-0.116		-0.101		-0.071		0.001		-0.039		-0.083	
	(0.207)		(0.180)		(0.155)		(0.040)		(0.064)		(0.066)	
Protective (Factor Score)	0.316		0.370	*	0.219		0.093	**	0.124		0.167	**
	(0.235)		(0.212)		(0.198)		(0.046)		(0.077)		(0.073)	
Social (Factor Score)	-0.216	***	-0.065		-0.040		-0.104	***	-0.102	***	-0.061	*
	(0.089)		(0.074)		(0.078)		(0.027)		(0.026)		(0.034)	
Understanding (Factor Score)	0.200	*	0.302	***	0.290	***	0.063		0.096	**	0.137	***
	(0.107)		(0.108)		(0.094)		(0.045)		(0.041)		(0.038)	
Values (Factor Score)	0.211		0.248		0.163		0.186	**	0.142		0.094	
	(0.208)		(0.161)		(0.161)		(0.079)		(0.088)		(0.072)	
Gender (Male)	-0.005		0.029		0.052		0.024		0.017		0.077	
	(0.149)		(0.136)		(0.117)		(0.050)		(0.066)		(0.059)	
Ln (Age)	0.083		-0.217		-0.005		-0.100		-0.173		-0.029	
	(0.235)		(0.231)		(0.313)		(0.080)		(0.116)		(0.126)	
Education (Self)	-0.085	***	-0.047		-0.039		-0.009		0.004		-0.014	
	(0.039)		(0.034)		(0.045)		(0.014)		(0.022)		(0.018)	
Ethnicity (Non-White)	-0.348		-0.317	**	-0.114		-0.073		-0.089		-0.076	
	(0.244)		(0.144)		(0.189)		(0.085)		(0.096)		(0.078)	
Relationship (Married/Relationship)	-0.211		-0.273		-0.306	*	0.001		-0.128	*	-0.229	***
	(0.140)		(0.186)		(0.166)		(0.075)		(0.066)		(0.054)	
Ln (Income)	0.024		0.023		-0.001		0.000		0.005		0.005	
	(0.017)		(0.018)		(0.021)		(0.005)		(0.009)		(0.008)	
Ln (Duration)	1.287	***	1.100	***	0.868	***	0.508	***	0.471	***	0.308	***
	(0.093)		(0.086)		(0.075)		(0.047)		(0.074)		(0.043)	
Constant Term	-4.638	***	-0.881		1.594		-2.517	***	-1.318	***	0.066	
	(1.128)		(1.103)		(1.177)		(0.345)		(0.480)		(0.550)	
Pseudo R-Squared	0.132		0.115		0.090		0.128		0.123		0.105	

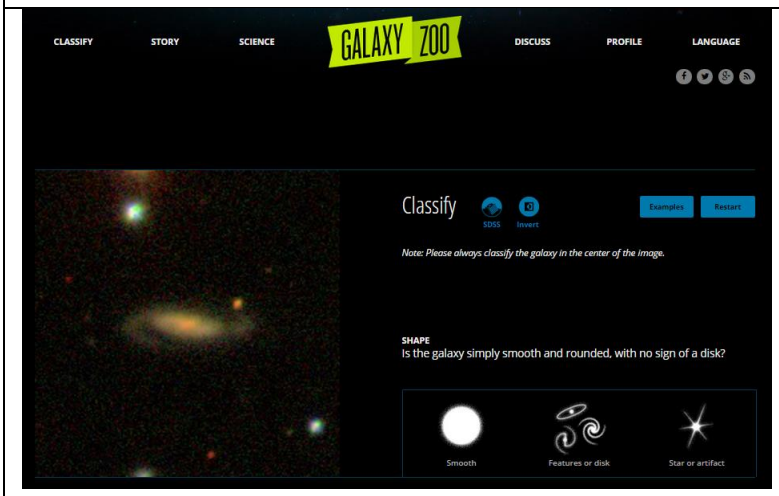
Significance: * = 90% level, ** = 95% level, *** = 99% level. Other controls included but not reported are as follows: Home Project Controls (Planet Hunters, Penguin Watch, Seafloor Explorer, Snapshot Serengeti, Residence (City), Education (Parents), Science Qualifications, Ln(Paid Work), Ln(Charity Donations), Religion, Number of Children (<12, <18, 18+).

Table 5b: Quantile Regression Analysis (Retention)

Engagement Measure	**(iii) Ln(Unique Days)**			**(iv) Ln(Active Period)**		
Quantile	**0.25**	**0.50**	**0.75**	**0.25**	**0.50**	**0.75**
Career (Factor Score)	-0.170 **	-0.251 ***	-0.269 **	-0.305 ***	-0.032	-0.011 *
	(0.069)	(0.056)	(0.109)	(0.116)	(0.030)	(0.005)
Enhancement (Factor Score)	0.021	-0.073	0.031	-0.092	0.021	0.003
	(0.101)	(0.087)	(0.141)	(0.173)	(0.051)	(0.010)
Protective (Factor Score)	0.069	0.185 *	0.139	0.195	-0.008	0.000
	(0.103)	(0.102)	(0.186)	(0.196)	(0.058)	(0.010)
Social (Factor Score)	-0.080	-0.062	-0.063	-0.093	-0.016	-0.009 *
	(0.060)	(0.044)	(0.068)	(0.075)	(0.016)	(0.004)
Understanding (Factor Score)	0.117 ***	0.144 *	0.093	0.064	0.026	0.010
	(0.065)	(0.082)	(0.079)	(0.082)	(0.023)	(0.006)
Values (Factor Score)	0.149	0.194	0.210	0.463 **	0.101 **	0.020 **
	(0.141)	(0.119)	(0.131)	(0.199)	(0.051)	(0.010)
Gender (Male)	-0.031	-0.116	0.014	0.021	0.031	0.003
	(0.085)	(0.095)	(0.108)	(0.116)	(0.033)	(0.007)
Ln (Age)	-0.104	0.042	-0.115	-0.044	-0.032	0.002
	(0.105)	(0.184)	(0.161)	(0.204)	(0.048)	(0.013)
Education (Self)	-0.029	-0.036	-0.003	-0.024	-0.016	-0.003
	(0.027)	(0.030)	(0.032)	(0.045)	(0.013)	(0.003)
Ethnicity (Non-White)	-0.104	-0.262 ***	-0.185	-0.203	-0.018	-0.006
	(0.105)	(0.097)	(0.182)	(0.290)	(0.040)	(0.011)
Relationship (Married/Relationship)	-0.049	-0.126	-0.160	-0.147	-0.039	-0.018 **
	(0.122)	(0.117)	(0.128)	(0.133)	(0.030)	(0.008)
Ln (Income)	0.021 **	0.005	-0.009	0.018	0.003	0.000
	(0.010)	(0.012)	(0.011)	(0.020)	(0.004)	(0.001)
Ln (Duration)	1.027 ***	0.940 ***	0.881 ***	2.625 ***	1.313 ***	1.105 ***
	(0.067)	(0.054)	(0.091)	(0.214)	(0.050)	(0.008)
Constant Term	-5.283 ***	-4.248 ***	-2.124	-13.218 ***	-2.048 ***	-0.839
	(0.755)	(0.731)	(0.797)	(1.630)	(0.361)	(0.069)
Pseudo R-Squared	0.164	0.132	0.120	0.284	0.283	0.360

Significance: * = 90% level, ** = 95% level, *** = 99% level. Other controls included but not reported are as follows: Home Project Controls (Planet Hunters, Penguin Watch, Seafloor Explorer, Snapshot Serengeti, Residence (City), Education (Parents), Science Qualifications, Ln(Paid Work), Ln(Charity Donations), Religion, Number of Children (<12, <18, 18+).

Figure 1: Volunteer Interfaces for Zooniverse Projects



(i) Galaxy Zoo

Launch:

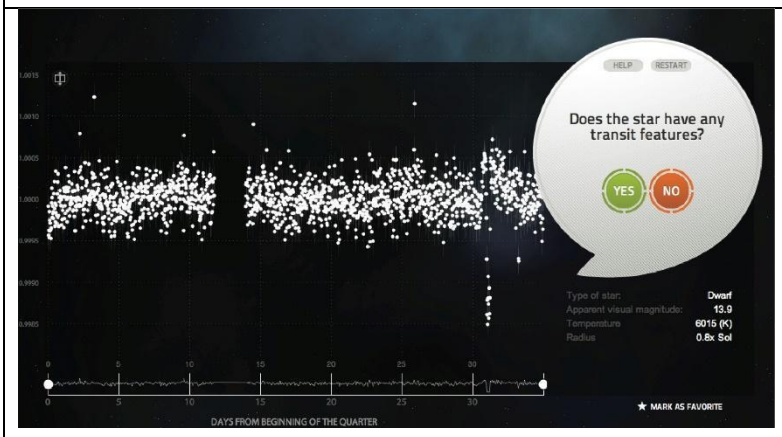
2007 (Galaxy Zoo 1); 2012 (Galaxy Zoo 4)

Number of Registered Volunteers

86,280 (Galaxy Zoo 4)

Brief Description of Task

Answer a series of questions relating to the shapes of deep space galaxies.



(ii) Planet Hunters

Launch

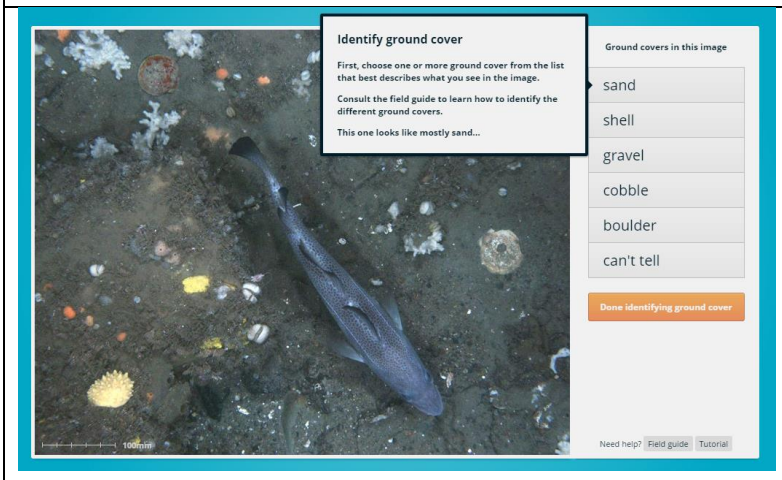
2010

Number of Registered Volunteers

172,628

Brief Description of Task

Identify drops in light that might indicate undiscovered planets passing in front of parent stars.



(iii) Seafloor Explorer

Launch

2012

Number of Registered Volunteers

21,508

Brief Description of Task

Indicate the type of ground cover and the presence, size and shape of marine life in photos of the sea bed.

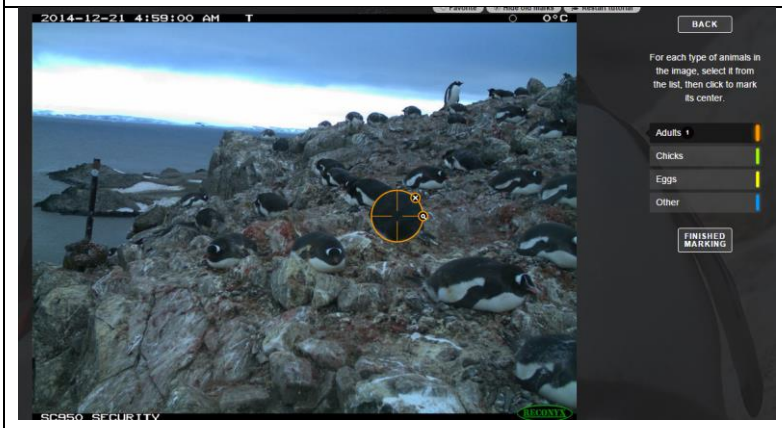
	**(iv) Snapshot Serengeti** Launch 2012 Number of Volunteers 32,429 Brief Description of Task Identify the number and types of animals appearing in images from camera traps on the Serengeti
	**(v) Penguin Watch** Launch 2014 Number of Volunteers 19,499 Brief Description of Task Mark the location and size of penguins appearing in images from the Antarctic.

Figure 2: Distribution of Classification Activity: Zooniverse Population versus Survey Sample