This project is co-financed by the European Union

Grant Agreement No.: 824603
Call: H2020-SwafS-2018-1
Type of action: RIA
Starting date: 1/02/2019

ACTION

# D 5.6 - Initial analysis and guidelines of incentives and motivation within citizen science

**Coordinator: Neal Reeves, KCL**
**Quality reviewer: Esteban González UPM**

| | |
|---|---|
| **Deliverable nature** | Report |
| **Dissemination level** | Public |
| **Work package and Task** | 5 |
| **Contractual delivery date** | M18 |
| **Actual delivery date** | M18 |

# Authors


| **Author name** | **Organization** | **E-Mail** |
|---|---|---|
| Neal Reeves | KCL | neal.reeves@kcl.ac.uk |
| Irene Celino | Cefriel | Irene.Celino@Cefriel.com |
| Gloria Re Calegari | Cefriel | Gloria.Re@Cefriel.com |

| | |
|---|---|
| **Abstract** | This deliverable sets out the initial analyses conducted by the ACTION team regarding the use and effectiveness of incentives, as well as distinct motivations and motivational factors within pollution citizen science. We analysed the motivations associated with participation in two pollution-related projects - one focused on light pollution and another on the impact of agricultural pollution on insect numbers. Additionally, we analysed the impact of financial incentives for incentivising participation in citizen science tasks, finding that while financial incentives can motivate participation, they are not associated with significant interest in tasks and may encourage low effort submissions and malicious behaviours. We note that these findings have significance for projects which may seek to combine paid crowdsourcing effort with volunteer generated data. We conclude with a set of preliminary guidelines for motivating and incentivising participation in citizen science, with a particular focus on pollution-related research. |
| **Keywords** | Motivations, incentives, survey, qualitative analysis, quantitative analysis |

**Disclaimer**





**How to quote this document**

*Reeves, N., Celino, I. and Calegari, G. (2020), Initial analysis and guidelines of incentives and motivations within citizen science*

# TABLE OF CONTENTS

# EXECUTIVE SUMMARY

This document presents analyses, case studies and set of guidelines for incentives and volunteer motivations within citizen science projects. We present background concepts and related work and build upon this existing literature to present current research within ACTION, partnering with our citizen science pilots to analyse contemporary motivations and the impacts of incentives on participation in pollution-related citizen science projects.

In this document, we first present a case study of motivations for participating in and networking with the TESS photometer network, a network of light pollution sensors disseminated and led by ACTION partner UCM. This case study combines and contrasts motivations and statistical analyses of responses from photometer holders with responses from crowdworkers in the Prolific platform. We found that photometer holders identify more with - and display stronger and more diverse motivations for - research and project activities than crowdworkers do for crowdsourcing tasks.

Following this, we explore the impact of payment-based incentives for longer-term engagement in citizen science activities. We set up the Cities At Night light pollution project within Amazon's Mechanical Turk platform and developed payment strategies to encourage medium- to long-term engagement. Our findings suggest that payments can motivate participation in such contexts, but such participation is associated with data quality issues and a lack of motivation and appreciation for scientific research.

Finally, we present a further case study of motivations for commencing and ceasing participation in a citizen science initiative run by ACTION partner DBC. Drawing on 2,455 records dating back to 1990, we summarise the most common reasons for volunteers beginning and ending their participation within the project, as well as the growth and changes in these motivations over time. Our findings suggest that while intrinsic motivations are most common, extrinsic needs -- such as administrative needs of the project -- are associated with significantly longer periods of participation overall.

We end this deliverable with guidelines for motivation in pollution citizen science and a conclusive summary of our findings.

# 1 INTRODUCTION

Citizen science projects rely on volunteer effort to gather and analyse data for scientific research purposes (Strasser et al., 2019). As shown by figure one, citizen engagement may be required at all stages of the research process and is therefore essential to the citizen science process. This raises an important question which we will investigate in this deliverable: what factor(s) motivate volunteer engagement within citizen science and how do different incentives impact these motivations and subsequent participation from volunteers?

In this deliverable, we present research findings from within ACTION, surrounding both incentives and motivations in pollution citizen science. Drawing on these findings and the background literature, we develop a set of guidelines for incentivising and motivating citizen engagement and understanding which motivations are central to volunteer participation and how these influence engagement.

This deliverable consists of the following parts:: in chapter 2, we present background literature and related work regarding motivations and incentives within citizen science. In chapter 3, we provide research from ACTION partner Cefriel exploring motivations for setting up light-pollution monitoring photometer sensors and comparing these to motivations within crowdsourcing more broadly. In chapter 4, we present research from ACTION partner KCL on the role and effectiveness of financial incentives within citizen science research and how the quantity of financial rewards influences medium- to longer-term engagement in citizen science research. Following this in chapter 5, we present a study of motivations for commencing and ceasing participation in a citizen science project run by Dutch Butterfly Conservation within ACTION. Within chapter 6, we present initial preliminary guidelines for motivations and incentives within pollution citizen science projects. Following this, we provide a brief of outline in chapter 7 of future work to be carried out prior to D 5.7, which will present the final set of guidelines on incentives and motivation within CS. We end this deliverable in chapter 8 with a concluding summary.

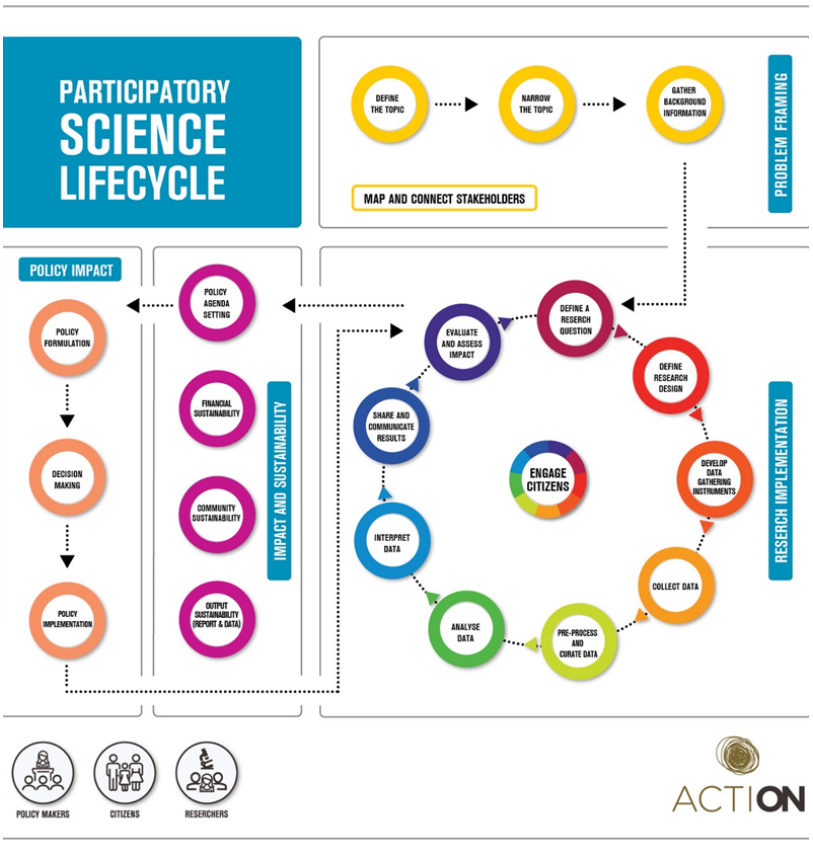


Figure One: Participatory Science Lifecycle Showing Stages of the Citizen Science Research Process

# 2 Background

## *2.1 Motivations*

Here we present three key forms of motivation that significantly impact participation and engagement within citizen science projects: intrinsic motivations, altruism and extrinsic motivations. Note that while altruism is itself a form of intrinsic motivation, we describe it distinctly because of its importance for long-term participation in citizen science and the rate at which altruistic motivations develop within CS projects, which are generally more slowly and at a somewhat later stage than other intrinsic motivations.

### *2.1.1 Intrinsic Motivations*

Intrinsic motivations refer to an individual's own personal interests and curiosity, without influence from external pressures, rewards and perspectives. In citizen science, these interests are often associated with the scientific research process, scientific field (e.g., climatology) or specific research topic (e.g., air pollution). Such motivation may not extend to the entirety of the scientific research process - participants do not necessarily display a heavy interest in how data are analysed and results drawn, instead being motivated more by the results being realised and disseminated (Aoki et al, 2017). Indeed, these motivations need not even stem from the scientific research process -- for example, in pollution monitoring projects, participants may be driven by their interest in their surroundings and local area, rather than the scientific research process itself (Maisonneuve et al, 2010). However, these appear to significantly underpin participation even in those projects with a significant number of extrinsic factors, such as heavily gamified projects and games with a purpose[1] (Curtis, 2015). These intrinsic factors may function as a feedback loop, where participants' intrinsic motivations and interests cause them to seek out opportunities to learn and explore these interests, which further strengthens intrinsic motivations (Jennett et al, 2016).

The importance of these intrinsic motivations have been highlighted across projects, studies and scientific fields. A study by Nov et al (2011) found intrinsic motives to be significant across projects -- particularly web-based projects -- and second only to collective motives stemming from the feeling of working together as a community for a common goal in driving participation within citizen science projects. More specific to pollution, a study of motivations in air pollution monitoring projects noted a variety of intrinsic factors associated with learning -- around health effects, the local area and a desire for increased knowledge for example -- as driving participation in such activities (Commodore, 2017). Moreover, environmental monitoring serves an important role in empowering citizens, with the desire to achieve goals -- communicating the dangers of pollution or curtailing environmental damage -- as a further motivator for engagement through citizen science (Aoki et al, 2017).

[1] Games which focus on "harnessing human skills for the purpose of research" where the game elements are "essential to motivate the public" (Lafourcade, Joubert and Le Brun, 2015).

Project administrators make use of these intrinsic motivations in a variety of ways. News and updates may be given on the scientific progress of the project, alongside updates on major discoveries and research both within the project and within the field more broadly (Schroer et al, 2018). At times, this may even overlap with extrinsically motivating factors. Scientific framing and information may also be used to shape and implement extrinsic factors such as games and competition campaigns (Schroer et al, 2018; Simperl et al, 2018). In some cases, particularly projects designed to run for a brief period of time, learning about the project discoveries has in itself been framed as a reward, with frequent updates from project scientists and opportunities to interact with them to learn more throughout the project (Reeves and Simperl, 2019).

### *2.1.2 Altruism*

Citizen science projects rely heavily on volunteer effort and the altruism of volunteers and this has been noted across various citizen science platforms. In the FoldIt Game With a Purpose, over 60% of respondents noted altruism and a desire to contribute to research as motivating factors for their participation within the project (Curtis, 2015), while a study of Zooniverse users noted that the desire to contribute to worthy causes and benefit society were the highest ranking motivational factors identified by volunteers, scoring 6.05 and 6.38 respectively on a 7-point likert scale. These altruistic motivations do not only concern society or science - a study of the Tomnod platform, which includes environmental and humanitarian CS activities found that the majority of respondents described a desire to improve the environment as a significant factor in their desire to participate in the platform (Baruch et al, 2016).

Nevertheless, there has been little direct analysis within the literature of altruism as a motivating factor for participation in and of itself. A 2012 study by Rotman et al found that while interest in projects, research and science are fundamentally important to drive *initial* participation in citizen science projects, it is altruism and a desire to assist science, further scientific research and work for the benefit of others that leads volunteers to contribute in the medium- to long-term. This finding has been reiterated in studies of the effectiveness of various motivational affordances in recruiting participants, which have demonstrated that altruistic framing is significantly less effective than framing around intrinsic and extrinsic factors (Lee et al., 2018).

### *2.1.3 Extrinsic Motivations*

Extrinsic motivations describe motivation arising from external factors that encourage completion of a given task, beyond or despite an individual's intrinsic motivations. These may include rewards, such as the provision of payments or prizes for engaging in a particular task, but also extend to more abstract factors such as feedback, the threat of a punishment or autonomy to make choices (Deci and Ryan, 2012). While citizen science is generally driven by *volunteer* effort, extrinsic factors have been trialled and implemented in a range of projects with varying degrees of success.

Most commonly within citizen science, extrinsic motivations are introduced through gamification, either as a fundamental element of a platform or as part of a temporary campaign or competition. These features may include features such as points, badges and achievements, leaderboards,

competitions, as well as prizes and even cash payments (Reeves et al, 2017). Beyond this, extrinsic motivations are rare in citizen science. While feedback may be offered within projects, such feedback is generally offered outside of task workflows, on a voluntary basis by fellow volunteers and scientists, with no associated benefits beyond the opportunity to learn associated with the feedback itself. As a result, feedback functions more as a social and potentially intrinsically motivating feature, rather than as an extrinsic factor (Jackson et al, 2016).

The potential negative impact of extrinsically motivating factors on an individual's intrinsic motivations are well documented across a range of contexts and this is no less true of citizen science (Kraut et al, 2012). Introducing fun, game-related elements has been noted to risk trivialising the important work being conducted and thus negatively influence volunteers' intrinsic motivations to participate (Ponti et al, 2018). While physical prizes and cash payments are uncommon in CS projects (Reeves et al, 2017), such rewards have been associated with undesirable and anti-social behaviours such as cheating and repeated, inaccurate submissions (Simperl et al, 2018). Additionally, pressures associated with competitions, rewards and associated behaviours have been identified as a source of potential stress that discourage more casual volunteers (Ponti et al, 2018).

## *2.2 Incentives*

Here we list some of the most commonly occurring incentives within the literature. It should be noted that this list is not exhaustive -- there are a number of project specific incentives and rewards, such as naming rights in projects that aim to discover new planets or asteroids. We instead focus on those incentives for which there is significant focus in the literature, or which have been associated with significant and clear outcomes in terms of motivation and engagement, through empirical studies and analyses.

### *2.2.1 Financial Incentives*

As noted, financial incentives are rare within citizen science. The EyeWire project offered cash prizes for a brief period as rewards for competition winners, but abandoned the approach after observing negative behaviours from participants (Simperl et al., 2018). Nevertheless, studies have experimented with offering financial incentives for citizen science activities. Mao et al. (2013) conducted a paid crowdsourcing experiment comparing engagement with the Galaxy Zoo project between volunteers and 3 groups of paid crowdworkers distinguished by high, medium and low per-task payment levels. While volunteers spent longer completing each task, paid crowdworkers completed significantly more tasks, regardless of their level of pay. Conversely, project administrators have identified tensions between paid and volunteer participants in citizen science, suggesting participants may be discouraged from participating where they are aware that others are doing similar work but receiving pay for doing so (Simperl et al, 2018; Woodcock et al, 2017).

### *2.2.2 Physical Incentives/Rewards*

Similarly, physical incentives and prizes are uncommon in citizen science projects, being more commonly offered than financial payments, but generally only on a temporary or infrequent basis and much more infrequently than more gamified features (Reeves et al., 2017; Reeves et al, 2018). Although these incentives serve predominantly as extrinsic motivations and rewards that motivate increased levels of engagement, notable examples have suggested a community-building element associated with these rewards, with participants driven to collect prizes as an example of their dedication to a project and to exchange these prizes with other participants, particularly newcomers (Simperl et al, 2018).

### *2.2.3 Gamification - points and badges*

Citizen science projects have adopted a number of gamification features to serve as incentives for participation, although the degree of integration of these features varies from project to project. In more gamified and competitive projects, participants may be ranked according to a leaderboard, with either the potential for rewards for those topping the leaderboard or a reputation mechanic where being at the top of the leaderboard is its own reward (Eveleigh et al., 2014). Leaderboards may capture only participation levels -- for example the number of contributions made by each participant -- or may use a more complex system such as a points system, overlapping with feedback by providing points for factors such as accuracy (Reeves, West and Simperl, 2018). A similar feature is the use of badges or achievements, with participants earning a virtual indicator of their achievements -- for example, upon completing 100 submissions -- which they can display within a project interface or on their social media page(s). As an additional motivator, these badges may be time-sensitive, overlapping with specific events such that participants can only earn a given badge by participating on a given day or week (Reeves et al, 2017).

### *2.2.4 Competitions*

An alternative form of incentivisation, competitions and temporary campaigns may be combined with the reward factors described above or as a one-off opportunity to drive participation. The EyeWire project offers regular competitions as weekly challenges and as seasonal narrative-driven events where volunteers participate within the project to drive the narrative forward, with both proving effective at driving participation from a significant proportion of the project (Reeves et al, 2018). The goals in such activities may drive participation between volunteers, or to encourage participants to work collaboratively to reach a common aim (Reeves et al 2018; Reeves and Simperl, 2019). The Loss of the Night ACTION pilot has carried out similar competitions, holding a raffle to allow contributors to the app to win a photometer.

# 3 Photometer and Prolific motivations

Studying motivation and investigating the factors influencing people's participation in citizen science projects is an essential aspect in the analysis of citizen science communities. Understanding the reasons that foster people to engage can support the successful design and implementation of effective participant involvement tasks, as well as pave the way for long-term engagement (Richter et al, 2018).

In this chapter we describe the methodology we adopted to study drivers of human behaviors inside communities of citizen scientists through surveys using the CONEY toolkit (explained in deliverable 4.1) and we present the results of two studies we conducted on two communities which differ in terms of both motivations and incentives.

## *3.1 Methodology*

The methodology adopted to investigate the motivating factors in citizen science communities through a survey can be summarized by the following list of steps and activities.

- **Survey preparation**: define the research question, define/select investigated factors, formulate questions to collect data for such factors, set-up the questionnaire in the tool by tagging questions with the respected factor and answers with their numerical coding, pre-test the survey with some users
- **Survey administration**: identify list of potential respondents, send survey to CS pilot participants, if necessary re-solicit responses
- **Collect & process survey responses**: monitor answer collection, export data, process data according to different statistical analyses, perform comparative analyses
- **Interpret and share survey results**: draw insights from result analysis, collect elements of the experiment, anonymizing data if needed, select suitable open licenses, openly publish research objects and results.

The next paragraphs explain in more details some of the main activities done in our investigation.

### *3.1.1 Research question definition*

To study the drivers of human behaviour we relied on existing questionnaires and surveys (Richter et al, 2018) used to evaluate the level of motivations of participants to Citizen Science projects. The reference methodology we started from is derived from the best practices from citizen science research (Levontin et al, 2018), which in turn are inspired from and extend the Schwartz Theory of Basic Values (Schwartz, 2012).

Based on findings from the literature, our research question is: which factors influence the motivation of CS participants?

### *3.1.2 Research design*

To answer the above research question, we designed a questionnaire as follows.

We selected 10 categories or latent variables mainly related to altruism and intrinsic motivation (*self-direction, stimulation, hedonism, achievement, power, conformity, benevolence, universalism, routine and belongingness*) that may be correlated to and may influence the main goal of the investigation, which is the global motivation of citizen scientists.

For each latent variable we defined 2 questions, obtaining a set of 20 questions in total. In addition to these questions for the investigation on the latent variables influencing motivation we also defined a question to directly ask participants to evaluate their level of motivation on a scale of 5-point likert scale. All the questions were designed to have closed answers: questions are annotated with the respective latent variable, while answers are associated with both a qualitative value (a textual label to be displayed to the respondent) and a quantitative value (the numerical coding for results analysis, ranging between 1 and 5).

The last question we added is an open question asking why participants decided to join the citizen science community. By letting people freely express their point of view, we would like to grasp the main motivations for participating.

These 22 questions represent the backbone of the survey we used to study motivation of participants in the two use cases explained hereafter. Actually this methodology is a conceptual framework that is not limited to this specific initiative but can be reused in different scenarios. By using the same annotations on questions and answers (i.e. the indication of the investigated latent variable for each question and the numerical coding for the answers) it is possible to easily and directly compare different survey result datasets.

### *3.1.3 Data gathering instrument*

The tool we chose for implementing the survey - Coney (explained in deliverable 4.1) - allowed us to add a "storytelling" component to the survey. This means defining the survey as a sequence of questions and conversational elements (text, images and gif) to make the survey more enjoyable and to help personalizing the conversation flow. In addition, different branching question and answer pairs customize the conversation according to the answer given by the user. The result is an interactive questionnaire that is experienced by respondents in a chat-like form.

### *3.1.4 Data collection*

The conversational survey was then administered to the target communities, as explained in the case studies below. When administering the survey through a crowdsourcing platform, the reliability of data collected is guaranteed by the insertion of a control question in the questionnaire to check the attention of the compilers while answering the survey. If the compilers answer this question in a wrong way we discard all the answers given by this user.

### *3.1.5 Data analysis and interpretation*

With the data collected we perform a quantitative analysis by averaging the value associated with answers in response to questions with the same latent variable tag. Then we correlate this value

with the level of global motivation indicated by the respondent. The goal is to study the relationships between motivating factors and global motivation in order to discover which motivating factors play a significant role in each specific scenario.

### *3.1.6 Results sharing and communication*

The complete dataset, including the survey structure, the collected answers and the analysis of correlation, is made openly available in RDF (modelled using the Survey Ontology https://w3id.org/survey-ontology) and CSV formats on Zenodo (Scandolari et al, 2020).

## *3.2 Case studies and results*

This section describes the two use cases analyzed, by detailing the context and the customization of the methodology to the specific scenario and by reporting the analysis done on the data collected.

### *3.2.1 Use Case 1: the TESS Photometer Network*

We study citizen scientists' motivation within the TESS network initiative[2], a citizen science community focused on fighting light pollution. The TESS network is composed of around 120 people hosting TESS photometers. Citizens are mainly involved in the data collection task, which consists of hosting and installing photometers to measure the level of sky brightness to fight light pollution.

We customized the formulation of the survey questions illustrated above in order to make them more specific to the TESS photometer context. We added 10 questions to further investigate the demographics of the volunteers, their level of engagement in the activity and how they use/plan to use the data collected.

We collected the answers from 83 volunteers, corresponding to 69% of our target users. This response rate is very successful compared to both the average survey response rate (33%) and the response rate of surveys sent by email (30%) (Lindemann, 2019).

As regards the custom questions, these are the results we collected. In terms of demographics, the 85% are male and, as shown in the pie chart below, the 70% of volunteers are older than 45.

[2] https://tess.stars4all.eu/network/

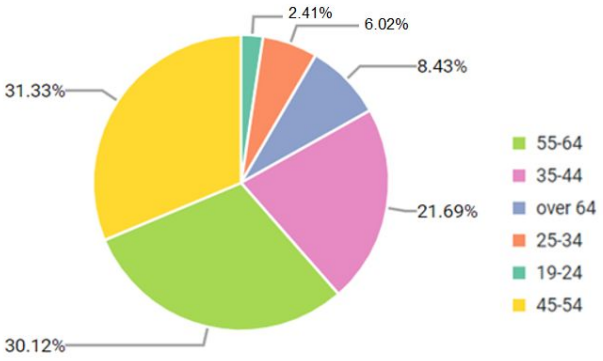


Figure Two: Pie chart showing responses to the question “how old are you?”

The volunteers are very diverse in terms of occupation: the majority of them defines themselves as amateur astronomers and light pollution fighters; ca. 30% of them are professional astronomers; about 20% work for astronomical outreach associations (museum, planetarium, dark sky association, etc.) or are astro-tourism actors.

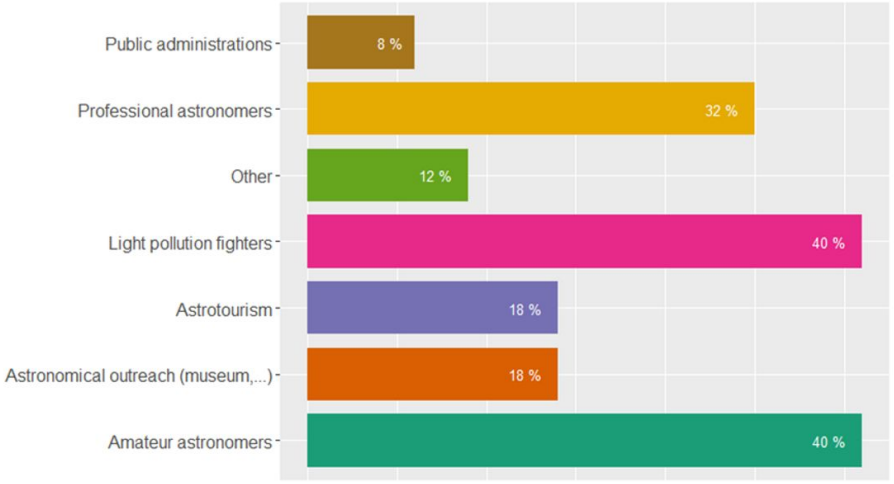


Figure Three: Bar chart showing responses to the question “Which of the following categories do you identify with most?” (a multiple choice question)

As regards the engagement in this specific CS project, we asked the participants how often they check their photometers on average. Only 18% never checked the photometer installed. This percentage demonstrates a high level of engagement in the majority of the volunteers.

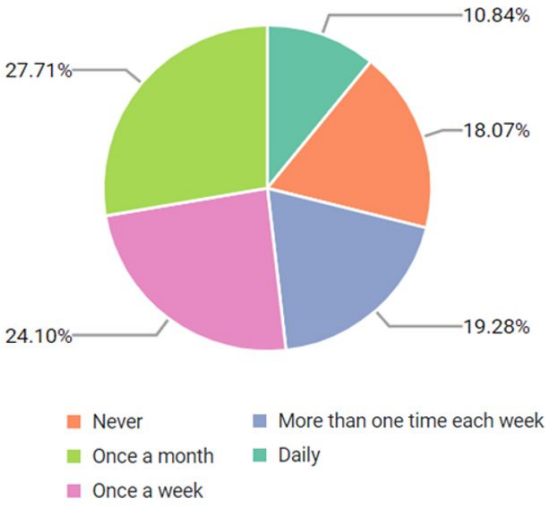


Figure Four: Pie chart showing responses to the question “How often do you check your photometer data on average?”

Another indicator of a high level of engagement and interest in the topic is that 75% of respondents are using or plan to use the data collected by their photometer and by the other photometers of the network. The figure below shows that 55 out of 83 volunteers are interested in downloading, exploring or using data collected from all the sensors installed.


Figure Five: Bar chart showing responses to the question “How much are you interested in downloading, exploring or using data collected from other photometers in the TESS network?”

Volunteers aim to use this data mainly for research (conference, journal, thesis...) and outreach activities about light pollution (talk, poster...). Some of them want to use data collected to obtain dark site certification for the region where they live.

As it happens with other environmental monitoring projects in the literature (cf. D5.1), TESS participants have the desire to achieve goals such as communicating the dangers of pollution or curtailing environmental damage and are very interested on their surroundings and local areas.

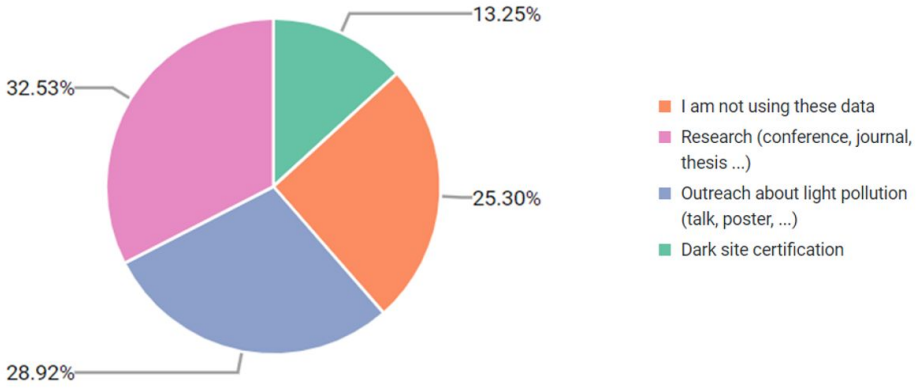


Figure Six: Pie chart showing responses to the question “what are you using (or planning to use) the data provided by your photometer for?

In the meantime, community members would like to know more about how the data collected are used by researchers. Actually, only 30% declares to have evidence about the usage of data collected. Increasing the sharing of the results of the experiments done can be an added value to help increasing the engagement level inside the community. Indeed, learning about the project discoveries has in itself been framed as a reward. This includes frequent updates from project scientists and opportunities to interact with them to learn more throughout the project.

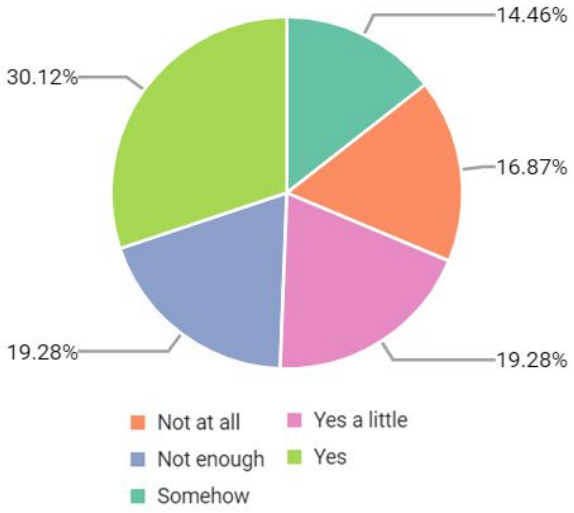


Figure Seven: Pie chart showing responses to the question “do you have evidence that data collected from the network has been used by researchers?”

The high level of interest in the TESS initiative is confirmed by the average value of global motivation that has been collected by asking directly to compilers their perceived level of motivation to participate in this initiative. The average value is 4.39 on a scale of 5 and the distribution of the values of global motivation is shown in the figure below.

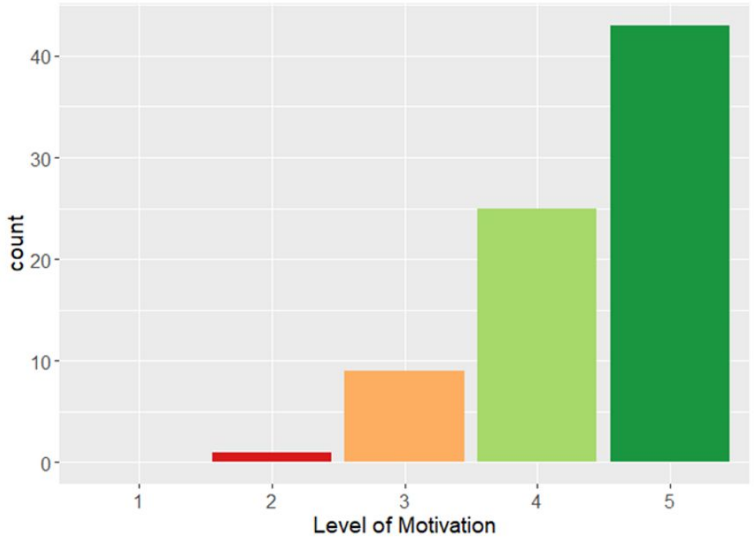


Figure Eight: Bar chart showing responses to the question “how much are you motivated in participating in the TESS network?”

As regards the Schwartz motivating factors, the table below shows the results of the survey: for each motivating factor the table reports the mean value of the answers and the correlation of each factor with the global motivation to participate. The level of significance of the correlation (p-value) is indicated by the stars next to the figures: *** p-value < 0.001, ** p-value <0.01, * p-value <0.05. The values of the answers range from 1 to 5 as well as the value of global motivation. The table also lists the questions asked for each motivating factor that have been properly customized for the specific scenario.

| Factor | Mean Answers | Correlation with global motivation | Questions |
|---|---|---|---|
| *Self-direction* | 4,43 | **0.491*** ** | How much do you expect to **learn** from your participation to the TESS network?<br>Are you **interested in topics** related to night sky brightness? |
| *Benevolence* | 4,42 | **0.62*** ** | How much do you see your participation in the TESS network as a **good thing to do**?<br>Do you participate to **contribute** and help the **scientific research**? |
| *Universalism* | 4,33 | **0.672*** ** | Do you participate for the possibility to make **data** about night sky brightness **more accessible**?<br>How much do you see your participation as a possibility to **raise public awareness** to the topic of this project? |

| | | | |
|---|---|---|---|
| *Hedonism* | 4,17 | **0.588*** ** | Does your participation to the TESS network make you **feel good about yourself**?<br>How **passionate** are you about the TESS network initiative? |
| *Stimulation* | 4,14 | **0.423*** ** | Did you join the TESS network to have the possibility to **do something new**?<br>Do you think your participation is an opportunity to **challenge yourself**? |
| *Achievement* | 4,13 | **0.424*** ** | Does the photometer represent an opportunity for you to **perform better than others** in some respects?<br>Does your participation to the TESS Network represent an opportunity to **do something meaningful**? |
| *Belongingness* | 3,75 | **0.456*** ** | Is your participation to the network influenced by the desire to **meet people** with **similar interests**?<br>By joining the TESS network, do you **feel part** of something worthwhile? |
| *Routine* | 3,08 | **0.272*** | Have you ever done night sky brightness measurement before (e.g. with other photometers)?<br>How **regularly** do you **participate** in citizen science projects? |
| *Power* | 2,83 | 0.156 | Do you believe you participation allows you to **gain recognition and status**?<br>Do you expect **something in return** from your participation to the TESS network? |
| *Conformity* | 2,35 | 0.075 | Do you **know other people** participating to the network?<br>To what degree were you **obliged** to participate? |

Table One: motivational factors, mean answer score, correlation with global motivations and associated questions for TESS phometer users

By looking at the mean values, we discover that participants expect to learn from their participation in the network and they are very interested in topics related to night sky brightness (4.43 of *Self-direction*). They want to participate because it is a good thing to do and because they want to contribute and help scientific research (4.42 of *Benevolence*). They are also very interested in participating to raise public awareness about the light pollution and to make data accessible for further researches (4.33 of *Universalism*). They are passionate about the TESS community and being part of this community make them feel better (4.17 of *Hedonism*).

The *Universalism*, the *Benevolence* and the *Hedonism* variables show also an high correlation with the global motivation (0.67 , 0.62 , 0.59 respectively), that means that people that shows a strong willingness in contributing to scientific research (raise awareness and make data accessible) and that are very passionate about the topic are also very motivated to participate to this citizen science community.

Other factors that push citizens to participate are the possibility to do something new, the chance to challenge themselves (4.14 of *Stimulation*) and the possibility to do something meaningful (4.13 of A*chievement*). People are not at all forced to participate and they participate even if they do not know anyone else of the community (2.35 of *Conformity*). They do not expect to have something in return and to gain recognition and status from the participation to the community (2.83 of *Power*)

Those results indicate very strong intrinsic motivations in TESS participants, that means that people participate because they are pushed by their own personal interests and curiosity, without influence from external pressures, rewards and perspectives.

As in the majority of citizen science communities, intrinsic motivation is associated with the specific scientific field and research topic (light pollution and measuring of sky darkness). Citizen scientists are not interested in all the steps of the scientific process but are more motivated by the results and their dissemination.

Intrinsic motivation is also associated with the altruism and volunteer effort of participants. Actually it is proved that what leads volunteers to contribute in the medium to long-term is the altruism and the desire to assist science, further scientific research and work for the benefit of others.

### *3.2.2 Use Case 2: the Prolific community*

The second scenario is about the study of the motivation of participants contributing to crowdsourcing activities in the Prolific platform[3]. Prolific is an on-demand, self-service data collection platform that helps in recruiting high quality research participants to take part in study, survey or experiment.

As in the previous use case, we customized the formulation of the questions in order to make them more specific to Prolific platform. In this case we were not interested in asking further questions to participants in addition to the ones related to the Schwartz variables. So at the end, the survey we administered was made up of 22 questions (2 global motivation + 20 Schwartz variables).

The average reward per hour is £7.85/hr and we estimated that the whole survey should take about 5 minutes to complete. Given this type of extrinsic reward, to prevent cheating and repeated, inaccurate submissions we discarded all the submissions that took less than 2 minutes to be completed.

We selected a significant sample of target respondents by filtering them by nationality (only people living in European countries), by minimum approval rate in past studies (80%) and by their past experience in crowdsourcing (at least 10 submissions in the platform).

[3] https://www.prolific.co/

We collected answers from 100 people, in order to have a comparable set of answers with respect to the TESS use case.
The average value of the global motivation is 3.58 (out of 5) with the distribution of values shown in the figure below. On average, people defined themselves quite motivated in participating in the Prolific platform.

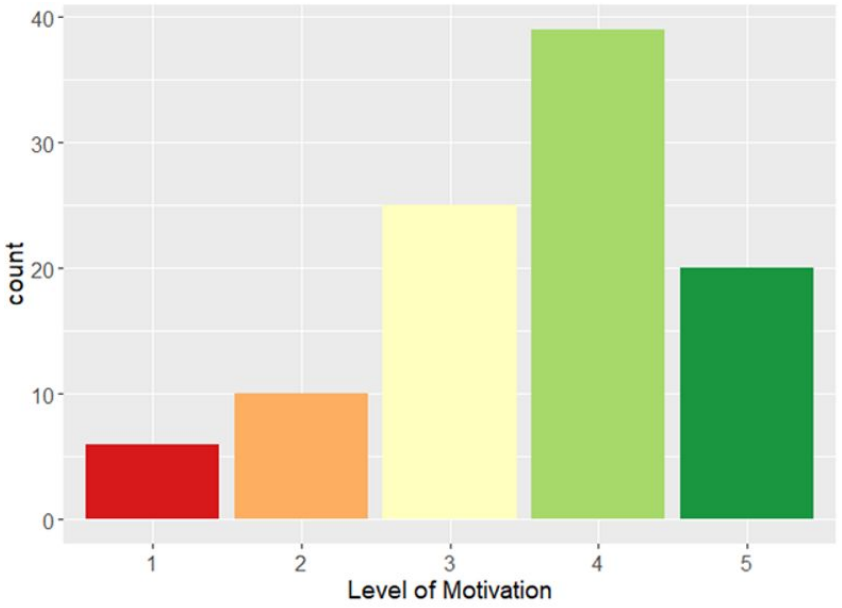


Figure Nine: Bar chart showing responses to the question “how much are you motivated in participating in the TESS network?” for Prolific workers

As regards the Schwartz motivating factors, the results obtained from the survey are listed in table below. For each each motivating factor the table reports the mean value of the answers and the correlation of each factor with the global motivation to participate. The level of significance of the correlation (p-value) is indicated by the stars next to the figures: *** p-value < 0.001, ** p-value <0.01, * p-value <0.05. The values of the answers range from 1 to 5 by design, as in the previous case study.

| Factor | Mean Answers | Correlation with global motivation | Questions |
|---|---|---|---|
| *Benevolence* | 4,00 | **0.717*** ** | How much do you see your participation in the crowdsourcing campaigns as a **good thing to do**? Do you participate to **contribute** and help the **scientific research**? |
| *Self-direction* | 3,82 | **0.623*** ** | How much do you expect to **learn** from your participation to crowdsourcing campaign? Are you **interested in crowdsourcing**? |

| | | | |
|---|---|---|---|
| *Stimulation* | 3,53 | **0.591*** ** | Did you join crowdsourcing campaigns to have the possibility **to do something new**?<br>Do you think your participation is an opportunity to **challenge yourself**? |
| *Achievement* | 3,53 | **0.499*** ** | Does the participation to crowdsourcing campaigns represent an opportunity for you to **perform better than others** in some respects?<br>Does your participation to crowdsourcing campaigns represent an opportunity to **do something meaningful**? |
| *Hedonism* | 3,39 | **0.556*** ** | Does your participation to crowdsourcing campaigns make you **feel good about yourself**?<br>How **passionate** are you about the crowdsourcing initiative? |
| *Belongingness* | 3,16 | **0.346*** ** | Is your participation to crowdsourcing campaigns influenced by the desire to **meet people** with **similar interests**?<br>By joining crowdsourcing campaigns, do you **feel part** of something worthwhile? |
| *Universalism* | 3,05 | **0.635*** ** | Do you participate for the possibility to make **data** about crowdsourcing campaigns **more accessible**?<br>How much do you see your participation as a possibility to **raise public awareness** to the topic of the crowdsourcing campaigns? |
| *Power* | 2,79 | **0.326*** ** | Do you believe you participation allows you to **gain recognition and status**?<br>Do you expect **something in return** from your participation to crowdsourcing campaigns? |
| *Routine* | 2,16 | **0.444*** ** | Have you ever done crowdsourcing campaigns **before**?<br>How **regularly** do you **participate** to crowdsourcing campaigns? |
| *Conformity* | 1,72 | 0.118 | Do you **know other people** participating to the crowdsourcing campaigns?<br>To what degree were you **obliged** to participate? |

Table Two: motivational factors, mean answer score, correlation with global motivations and associated questions for Prolific workers

By looking at the mean values, we discover that people participate in crowdsourcing campaigns because they consider it a good thing to do and because they want to contribute and help scientific research (4.00 of *Benevolence*). They expect to learn from their participation to crowdsourcing campaigns and they are interested in crowdsourcing in general (*Self-direction* 3.82).
People are not at all forced to participate and they participate even if they do not know anyone else of the community (1.72 of *Conformity*)

The value 2.79 of the *Power* factor indicates that *s*ome people may participate because they expect something in return (the provision of payment for the task done), indicating the presence of extrinsic motivation in some of the participants. Actually extrinsic motivations describe motivations arising from external factors that encourage completion of a given task.

On average people do not participate regularly and they have few experience in participating to crowdsourcing campaign (2.16 of *Routine*). This could be a consequence of the presence of monetary incentives that in the long-term are less effective than intrinsic motivations. This can generate a high turnover within the members of the Prolific community.

The motivating factors that most correlates with global motivation are *Benevolence*, *Universalism, Self-direction* and *Stimulation*. This means that the values of each of these motivating factors and the values of global motivation have the same statistical distribution (high values for both the global motivation and the motivating factor and low values for both the global motivation and the motivating factor). Really motivated people are the ones that want to help the scientific research and that want to make data more accessible. They are also interested in learning and would like to do something new that challenge themselves. This reflects the nature of this platform, that mainly recruits researchers and graduated participants with an high interest in science. The desire of learning new thing while contributing to science is a driver for participation already known in the literature (Commodore, 2017). All these factors are linked to intrinsic motivations and this demonstrates that what make feel people really motivated are topics and activities linked to their own personal interest and curiosity.

### *3.2.3 Comparison between the two use cases*

The two scenarios selected are quite different both in terms of motivations and incentives.

In the first case participants do not receive any monetary rewards and the only tangible return they can have is the data collected by their sensors. On the other hand, in the second scenario people receive a financial return when they complete the data collection task.
In Prolific respondents are called to contribute to different data collection tasks about very diverse topics. Conversely, TESS is a community focused on light pollution issues and the data collection tasks are related only to this topic. People in TESS are pushed only by intrinsic motivation whereas in the Prolific platform there is a mix of intrinsic and extrinsic ones.

By looking at the values of global motivation reported by compilers, we noticed that participants to TESS network are more motivated than participants of crowdsourcing campaigns, actually the average values are respectively 4.39 and 3.58. Also the distributions of values are different, with the majority of participants of TESS voting 5 and the Prolific ones voting 4. In addition, in the Prolific case, the 6% revealed to be not at all motivated to participate.

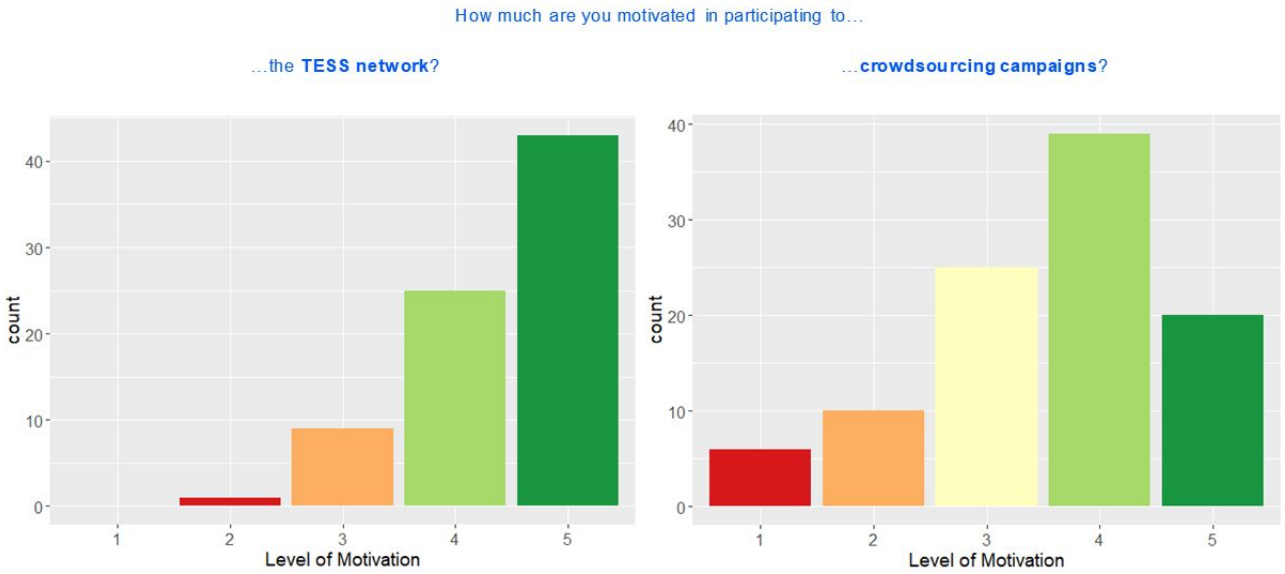


Figure ten: comparative motivations for TESS photometer holders and Prolific workers

The motivations that drive the two communities are also different and are highlighted in the word clouds below, which has been created starting from the open question about the main motivations. In the TESS case, the motivations are linked to the specific issue of light pollution, sky darkness/brightness measurement and to the possibility of collecting new data for research. On the other hand, participants of the Prolific platforms aims to help the scientific research as well but they also participate for money.

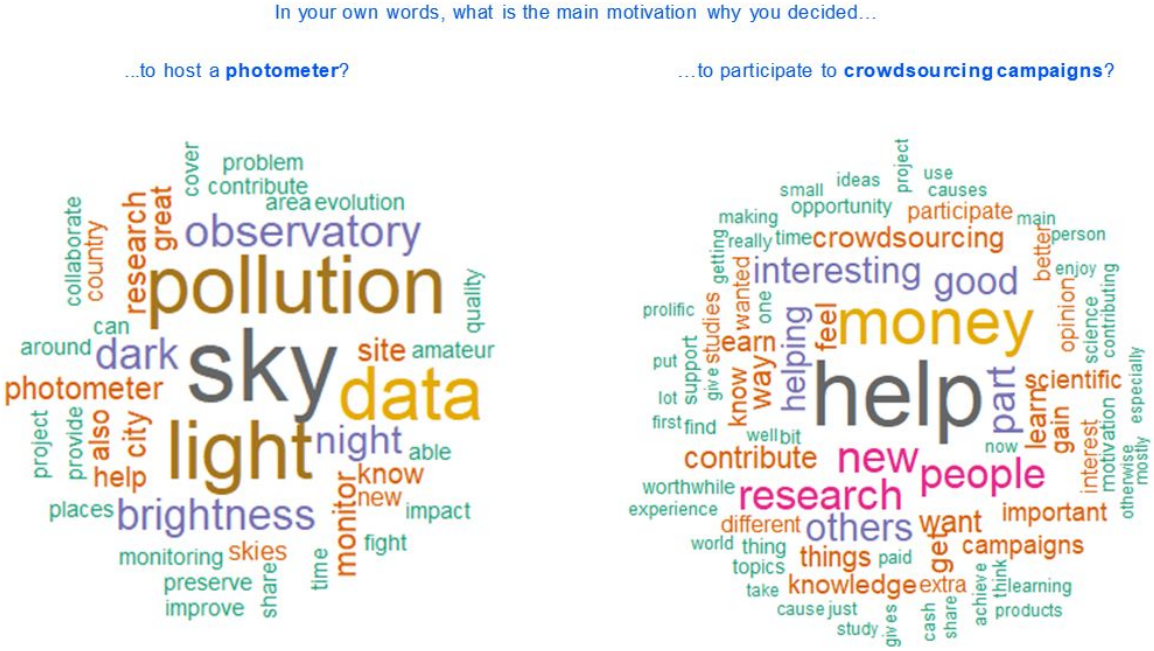


Figure eleven: comparative word clouds for TESS photometer holders and Prolific workers

By analysing the results of Schwartz variables, we discovered both commonalities and differences in the means and correlations between latent variables in the two use cases.

By comparing the mean values, it is evident that all motivating factors of TESS use case on average are higher than the same factors in the Crowdsourcing one. In both cases *Conformity* has low mean values indicating that in both cases people are not obliged to participate and they enter the community even if they do not know anyone else there. On the other hand, *Benevolence and Self-direction* have the highest variables in both cases, indicating in both cases an high interest in helping scientific research and in learning new things.

| Factor | Mean TESS | Mean Prolific | Questions |
|---|---|---|---|
| *Self-direction* | 4,43 | 3,82 | want to learn<br>interested in topics |
| *Benevolence* | 4,42 | 4,00 | good thing to do<br>contribute and help the scientific research |
| *Universalism* | 4,33 | 3,05 | making data more accessible<br>possibility to raise public awareness |
| *Hedonism* | 4,17 | 3,39 | making you feel good about yourself<br>how passionate are you |
| *Stimulation* | 4,14 | 3,53 | possibility to do something new<br>to challenge yourself |
| *Achievement* | 4,13 | 3,53 | perform better than others<br>do something meaningful |
| *Belongingness* | 3,75 | 3,16 | meeting people with similar interests<br>feeling part of something worthwhile |
| *Routine* | 3,08 | 2,16 | task already done before<br>frequency of participation |
| *Power* | 2,83 | 2,79 | gain recognition and status<br>expect something in return |
| *Conformity* | 2,35 | 1,72 | know other people<br>obliged to participate |
| ***Global Motivation*** | **4,39** | **3,58** | |

Table three: comparative global motivation factors, mean scores and questions for TESS photometer holders and Prolific users

Regarding correlations, the *Universalism* latent variable highly correlates with the global motivation in both cases (0.672*** in *TESS* and 0.635*** in *Prolific*) as well as the *Benevolence* variable (0.62*** in *TESS* and 0.717*** in *Prolific*). This means that in both cases really motivated people

decided to enter the community to make data more accessible, to raise public awareness and to contribute to scientific research. This confirms that the participants of both communities are very keen on science and research and are pushed by intrinsic motivations.

On the other hand, there are three correlations not present in the *TESS* but very significant in *Prolific*. One is between the motivation and the possibility to do something new and the opportunity to challenge themselves (0.591*** for S*timulation* ); the second one is the correlation between the motivation and the desire of learning something (0.623*** for *Self-direction).* Actually very diverse data collection tasks are assigned to the workers that can challenge themselves every time with a task about a new topic. The novelty and the diversity of tasks is the driver that pushes them to participate. The third one is about the *Power* variable, that may indicate that *TESS* participants are more selfless than *Prolific* ones because they are less interested in a financial return or physical rewards.

| Factor | TESS | Prolific | Questions |
|---|---|---|---|
| *Achievement* | **0.424*** ** | **0.499*** ** | perform better than others<br>do something meaningful |
| *Belongingness* | **0.456*** ** | **0.346*** ** | meeting people with similar interests<br>feeling part of something worthwhile |
| *Benevolence* | **0.62*** ** | **0.717*** ** | good thing to do<br>contribute and help the scientific research |
| *Conformity* | 0.075 | 0.118 | know other people<br>obliged to participate |
| *Hedonism* | **0.588*** ** | **0.556*** ** | making you feel good about yourself<br>how passionate are you |
| *Power* | 0.156 | **0.326*** ** | gain recognition and status<br>expect something in return |
| *Routine* | **0.272*** | **0.444*** ** | task already done before<br>frequency of participation |
| *Self-direction* | **0.491*** ** | **0.623*** ** | want to learn<br>interested in topics |
| *Stimulation* | **0.423*** ** | **0.591*** ** | possibility to do something new<br>to challenge yourself |
| *Universalism* | **0.672*** ** | **0.635*** ** | making data more accessible<br>possibility to raise public awareness |

Table four: comparative global motivations, mean scores and questions for TESS photometer holders and Prolific workers

# 4 Long-term Financial Incentives Crowdsourcing (paid citizen science activity)

Although citizen science projects are generally reliant on altruism and volunteer contributions, projects have experimented with using physical prizes and even cash payments as rewards for participation (Reeves et al, 2017; Simperl et al, 2018). Projects may have limited resources to offer participants and citizen science can often be seen as a low-cost way to gather large volumes of data, particularly when compared to full-time workers (Buytaert et al, 2014; Reeves, West and Simperl, 2017; Tweddle et al, 2012). Nevertheless, projects such as EyeWire and Zooniverse have previously experimented with cash payments, particularly during brief campaigns (Simperl et al, 2018; Mao et al, 2013).

Research into the effectiveness of these rewards has been mixed. While more web-based citizen science initiatives have been seen as a form of 'crowdsourced science' or crowdsourcing applied to science (Wiggins and Crowston, 2011), citizen science is generally distinct due to its focus on volunteer effort and intrinsic -- rather than extrinsic -- motivations. What little experimentation has been conducted in this area has used one off tasks to understand crowdsourcing engagement (Mao et al, 2013). But citizen science initiatives are highly reliant on *continued* participation from those volunteers who return to complete tasks on multiple occasions (Ponciano et al, 2014; Simperl et al, 2018).

In this chapter, we present an analysis of participation in citizen science tasks across a period of 10 days, using varying financial incentives to drive participation rather than relying on altruism and intrinsic motivations. We analyse daily participation, task abandonment rates, task accuracy and the occurrence of low-effort submissions, to monitor how participation varies across the task. We further present the results of a follow-up survey intended to further explore the motivations for participating in the task.

## *4.1 Methodology*

For this experiment, we used a pollution-specific citizen science task, drawn from the light pollution project *Cities At Night*, originally run by ACTION partner UCM. The second was a more generic crowdsourcing task, in the context of disaster relief, with less clear grounding in a specific research issue or question.

### *4.1.1 Task One*

Dark Skies ISS is a task from the Cities At Night[4] light pollution related citizen science project, which asks participants to classify images taken by astronauts from the International Space Station according to the content of the image. These images are then used to observe and monitor how

[4] https://citiesatnight.org/

light polluted different cities and locations are. The task itself is relatively simple and asks participants to look at an image, before selecting one of six categories: city, aurora, astronauts, stars, black/blank images or empty (no) image.

To focus on financial incentives and remove other incentives that may affect the levels of participation within the task, we recreated the cities at night task in a simple, survey-driven interface which can be seen in figure eleven below. This covered one main question, with two branching subquestions dependent on the response:

1. What is the content of this image?
    a. Astronaut
    b. Aurora
    c. City
    d. Stars
    e. Black
    f. No image
2. What is the quality of the image?
3. How clear is the image?

Questions 2 and 3 were triggered only in the event that participants selected ‘city’ for the first question, in accordance with the task design of the original cities at night task.

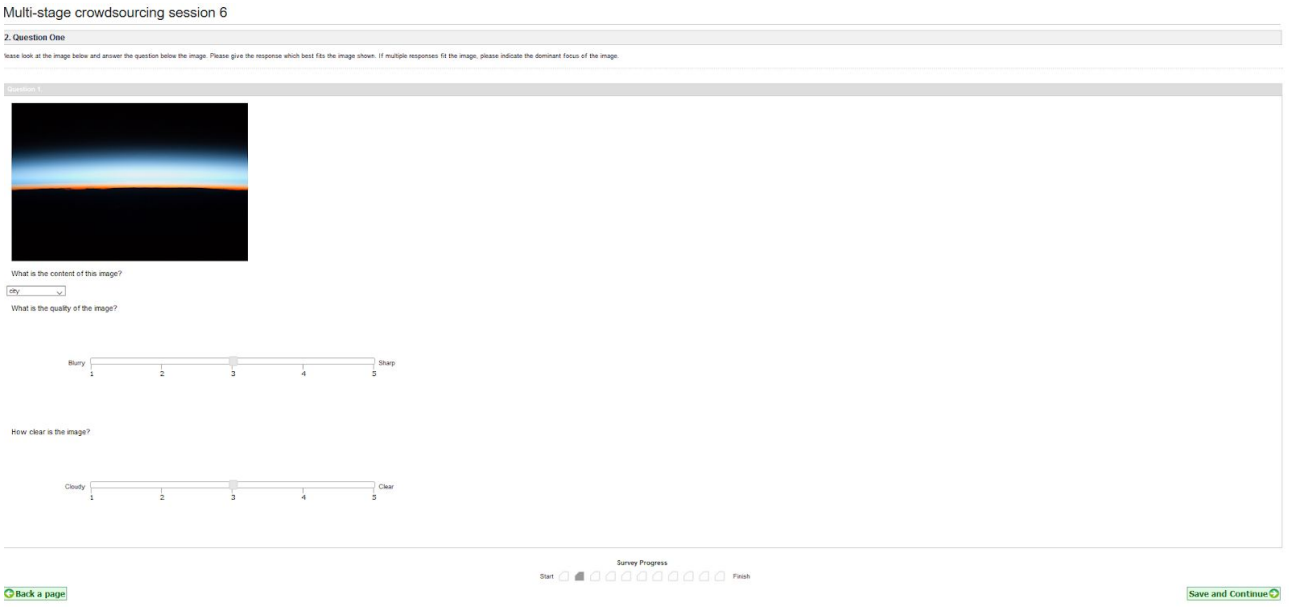


Figure eleven: task interface for the first citizen science experiment based on the Cities At Night project.

### *4.1.2 Task Two*

As a comparison with and extension of the findings of task one, we implemented a more generic crowdsourcing task, for which the activities were not clearly related to the task presented to volunteers. This task presented tweets pertaining to the aftermath of natural disasters - specifically hurricanes Harvey, Irma and Maria - and asked the volunteers to perform **Named Entity Recognition** tasks, picking out names, locations and occupations within the tweets and give three labels for each tweet. The task interface can be seen in figure twelve below.

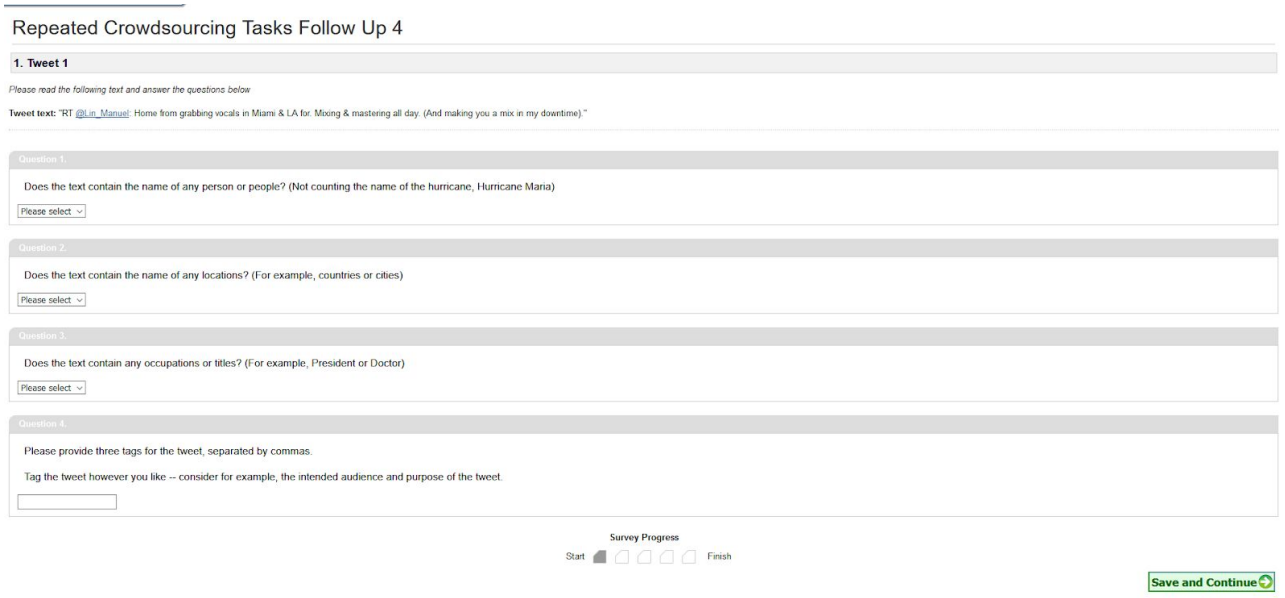


Figure twelve: task interface for the generic crowdsourcing task using disaster relief tweets.

### *4.1.3 Incentives and Rewards*

For each experiment, we assigned participants at random to one of three experimental conditions, such that each condition consisted of an equal number of participants. These experimental conditions were otherwise identical and varied solely on the basis of the financial incentives offered per chance. These conditions consisted of a low payment condition offering below average hourly earnings, a medium condition offering roughly average hourly earnings and a high payment condition offering above average hourly earnings. These hourly earnings were calculated on the basis of average hourly earnings as identified by Hara et al. in 2018. Due to the large variation in average earnings and the general degree of uncertainty surrounding average earnings, we set the *low* earnings at just below the low average calculated by Hara et al (2018) and the medium earnings at the high average calculated by Hara et al (2018). The high earnings were then equally spaced above the high average, such that there was an equal gap between the low and medium and medium and high rewards.

Although the two experiments were calculated on the basis of average earnings, the specific figures varied based on task. This was based on pilot studies which demonstrated that the two tasks took significantly different times to complete. As a result, while the average earnings for the

three groups were similar in both experiments, the specific rewards were different. These values were:

| Condition | Experiment One | Earnings experiment one | Experiment Two | Earnings experiment two |
|---|---|---|---|---|
| **Low** | $0.30 | $4.50 | $0.80 | $4.80 |
| **Medium** | $0.40 | $6.00 | $1.05 | $6.30 |
| **High** | $0.50 | $7.50 | $1.30 | $7.80 |

Table five: experimental conditions and associated per-task rewards and hourly earnings for each experiment

### *4.1.4 Experiment Platform*

Each experiment was run through Amazon's paid microtask crowdsourcing platform Mechanical Turk[5]. Tasks were launched through the TurkPrime (now Cloud Research) service, which allowed tasks to be queued in advance of the launch, with the opportunity to contact participants to send task reminders and easy facilities to assign qualifications to workers. These qualifications were used to ensure that only participants from the assigned conditions could see and access the tasks.

Mechanical Turk was chosen in part due to being one of the largest crowdsourcing platforms and having generally higher earnings than many other platforms, which are associated with often very low and very unethical payments for participants (Berg et al., 2018). While there are a number of citizen science platforms through which such a task could be launched, none allowed for direct payments to volunteers and we felt that launching a separate service to run the experiment would run the risk of biasing results due to volunteers' lack of trust in an unproven crowdsourcing platform.

### *4.1.5 Experiment Design*

Each experiment was conducted over a series of 10 experiments, launched over a 10 day period. We first launched a recruitment task, intended to recruit participants and allow for quality assurance, to the Mechanical Turk site through the TurkPrime service. This recruitment task asked volunteers to read a participant information sheet and offer their informed consent to take part in the experiment. The recruitment task also included a set of task questions and attention check questions designed to ensure participants were answering questions correctly. All participants completing this task were awarded the low payment value.

[5] http://mturk.com/

Participants who completed the recruitment task correctly were then given the opportunity to complete a set of nine follow-up tasks. These tasks were launched over nine consecutive days, with each available for just 24 hours. Participants did not have to complete any given task to participate in the next task -- that is, if a participant missed a day, he/she could complete the next day's task. For each daily task, three identical versions were launched, which varied only according to the level of pay offered: one with a low, one with a medium and one with a high reward respectively. Each daily task could only be seen by participants assigned to that particular payment condition. New participants who had not completed the recruitment task and been assigned to a group and existing participants who had been assigned to a different group could not see tasks other than their own.

For each daily task, we monitored the number of participants who successfully completed the task. Additionally, we monitored the number of participants who started the task, but abandoned it before completion. Following the closure of the experiment, we calculated the number of tasks completed by each individual worker, the accuracy of each worker and the amount of time taken by each worker to complete the daily task.

### *4.1.6 Follow Up Survey*

To better understand the motivations of individual participants, we launched a follow-up survey after each respective experiment was completed. This survey consisted of five statements, to which participants were asked to state how much they agreed based on a five-point likert scale, where 5 was strong agreement and 1 was strong disagreement. Additionally, participants were given the opportunity to identify factors that motivated their participation in the task and factors which motivated their abandonment of tasks using free text responses.

### *4.1.7 Data Analysis*

For each experiment, we compared participation between the three experimental conditions using statistical tests. Initially, normality tests were conducted to identify whether parametric or non-parametric tests should be used. Where the conditions for a parametric test were met, we used the **Analysis of Variance (ANOVA)** test, while for those cases where the conditions were not met (e.g., where values were not normally distributed), we conducted a Kruskal-Wallis H test. Where these tests detected a statistically significant difference, we then conducted further pair-wise testing to identify differences between groups.

## *4.2 Results*

### *4.2.1 Tasks per worker*

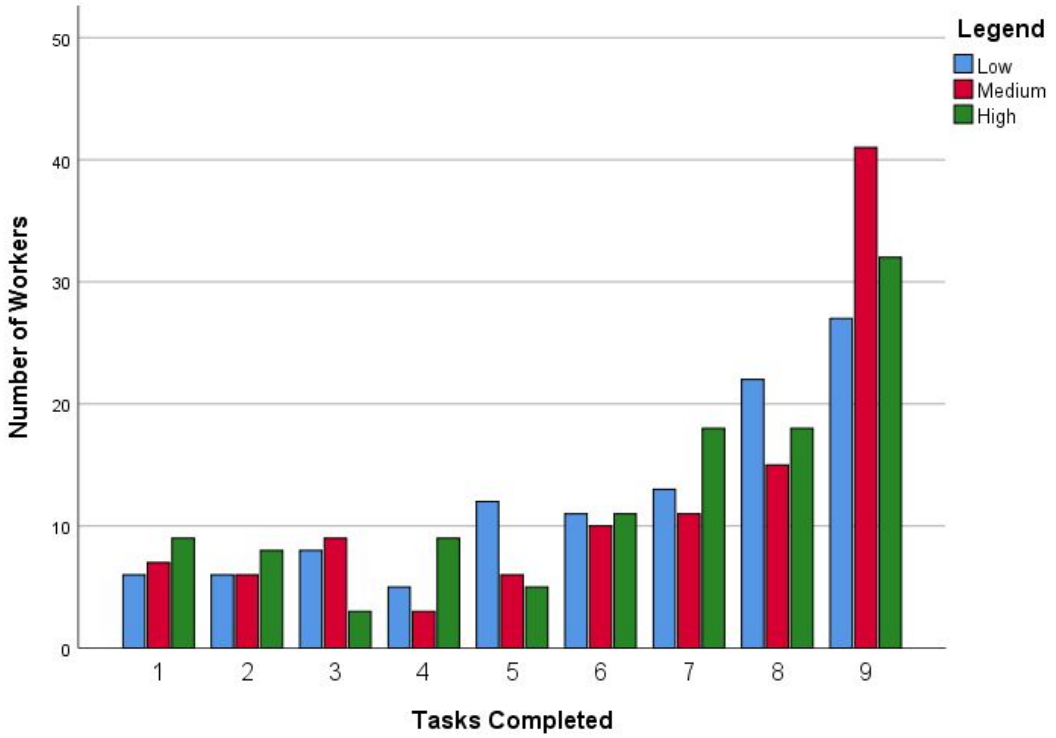


Figure thirteen: tasks completed per worker from low, medium and high payment conditions in experiment one

The majority of participants in both experiments completed 8 or 9 of the follow-up tasks. Although Figure thirteen shows that participants from the medium condition were most likely to complete all nine tasks, an ANOVA test showed **no statistically significant difference** ($p = 0.30$) between conditions in experiment one. In experiment two, using a text-driven and more generic research task, a non-parametric Kruskal Wallis H test initially suggested a statistically significant difference between groups ($H=10.60$, $p=0.005$). Ultimately, however, a follow-up Dunn's test noted a statistically significant difference between the medium and high payment conditions only ($Z=-3.14$, $p = 0.005$). The daily task completion numbers can be seen in figure thirteen and figure fourteen for the first and second experiment respectively.

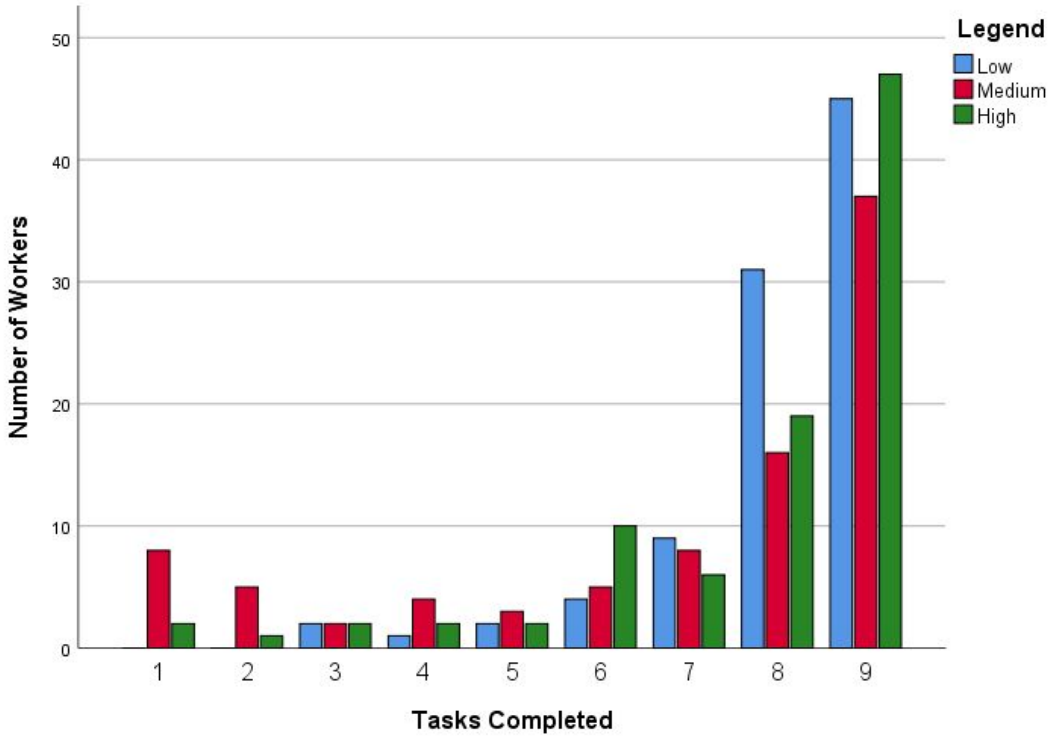


Figure fourteen: tasks completed per worker from low, medium and high payment conditions in experiment two

### *4.2.2 Daily Worker Numbers*

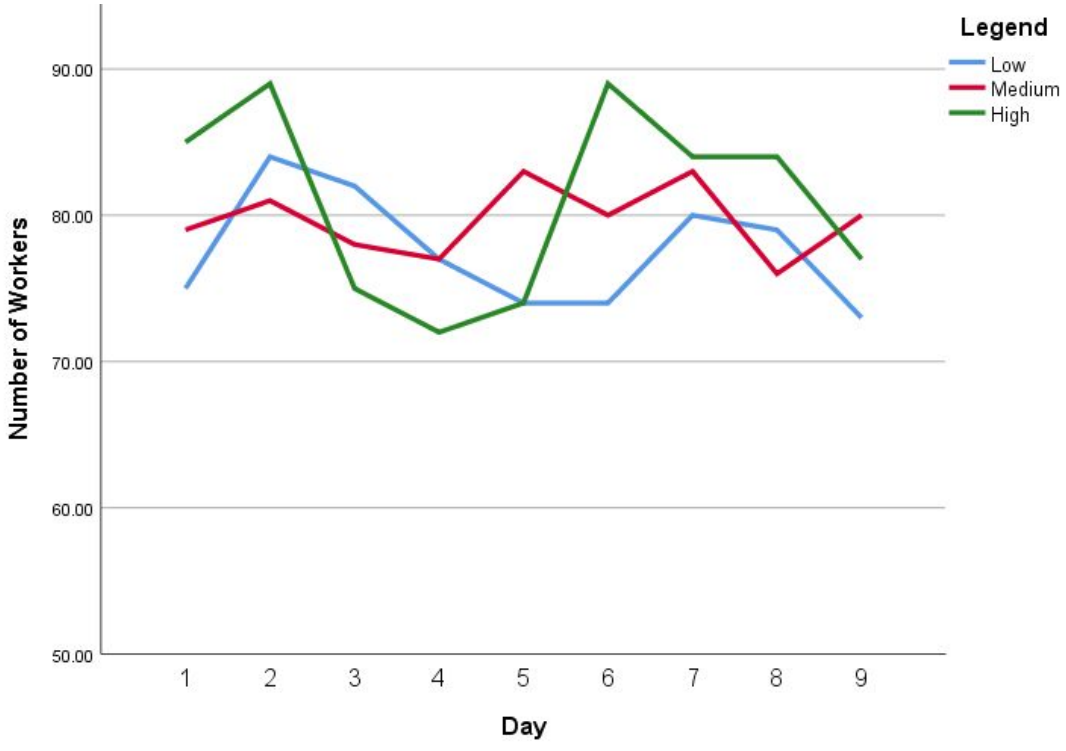

Figure fifteen: number of daily workers for low, medium and high payment condition over nine daily tasks for experiment one

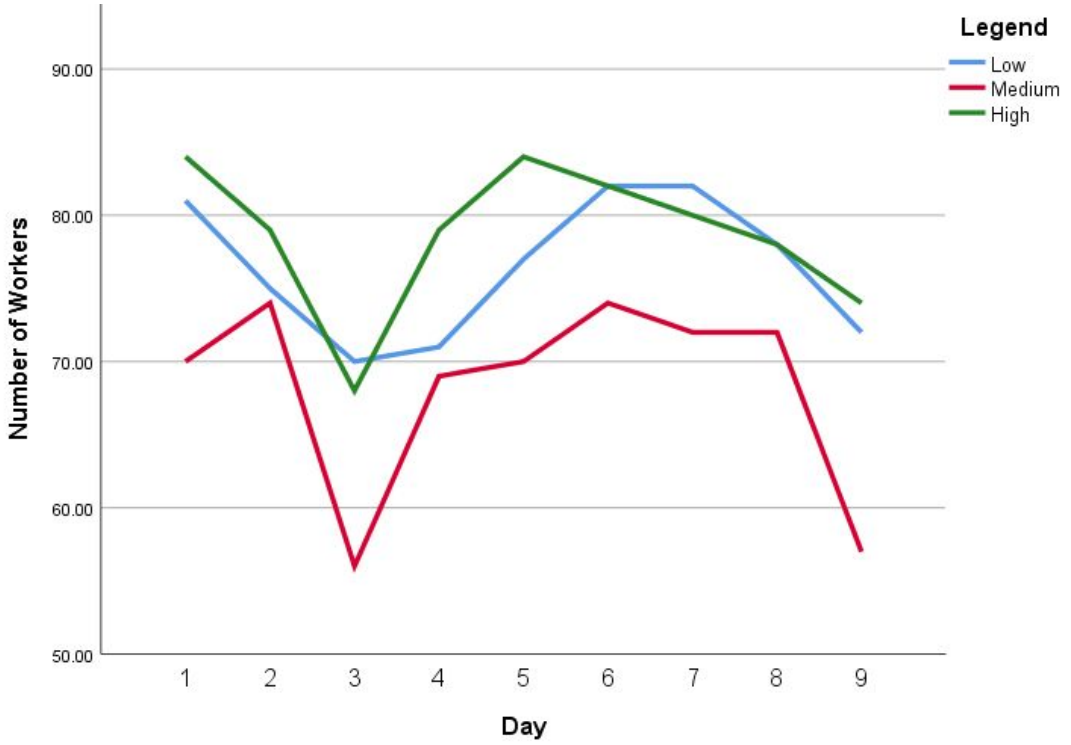


Figure sixteen: number of daily workers for low, medium and high payment condition over nine daily tasks

Figures fifteen and sixteen show daily worker numbers for the three payment conditions for each of the two experiments. An ANOVA test of daily worker counts for experiment one found no statistically significant difference between the three groups (p=0.30). A Kruskal-Wallis H test for experiment two initially found a statistically significant difference between groups (p=0.005). The follow-up Dunn's test revealed a statistically significant difference between the medium and high payment condition only (p=0.005).

### *4.2.4 Task Abandonment*

Each worker was provided with a unique URL for each daily task and we monitored whether these URLs were used, such that we could identify whether an individual worker had attempted to complete the task. We monitored the number of workers for each group that followed these URLs but did not submit a completed task as a means to monitor task abandonment within the two experiments. Mechanical Turk does not otherwise other a means to identify abandoned tasks and TurkPrime obfuscates this data by offering a 'bounce rate' which does not offer a full breakdown of abandonment figures.

An ANOVA test for the image-driven citizen science task found **no statistically significant difference** between the three groups (p=0.94). Similarly, an ANOVA test for the second text-driven task found **no statistically significant difference** between the three groups (p=0.71).

### *4.2.5 Accuracy*

Although participants had to complete the recruitment task accurately to begin the follow-up tasks in both experiments, this was no guarantee that participants would complete the follow-up tasks accurately. To assess the accuracy of volunteer submissions, an expert classification was made as a gold standard for each of the 90 images and 45 tweets. Each volunteers' submission was then compared with this gold standard and assigned a score of 1 (accurate) or 0 (inaccurate). Accuracy scores were then aggregated for each question according to experiment and payment condition. There was significant variation in aggregated accuracy within and between payment groups across both experiments.

We conducted a Kruskal-Wallis H test for each experiment comparing these aggregated accuracy scores between the three groups. The results demonstrated **no statistically significant difference** for either the first (p=0.83) or the second (p=0.89) experiment.

### *4.2.5 Low effort and malicious submissions*

Across each of the experiments and payment conditions, we received submissions from workers who had failed to complete tasks correctly, either ignoring instructions, or making malicious submissions which aimed to circumvent any quality assurance processes in place. These included selecting responses to avoid branching questions and copying and pasting irrelevant or identical text in response to some or all questions.

A Kruskal-Wallis H test analysis of these low effort submissions between groups in experiment one found a statistically significant difference between groups (H = 12.72, p = 0.002). Follow up Dunn's tests showed a statistically significant difference between the medium and high (p=0.003) and medium and low conditions (p=0.015). Similarly in experiment two, a Kruskal-Wallis H test found a statistically significant difference between groups (H = 11.95, p = 0.003). The subsequent Dunn's test found only a statistically significant difference between the low and high payment condition (Z = -3.46, p = 002).

### *4.2.6 Survey Responses - Likert Scores*

| Motive | 1 - Low | 1 - Medium | 1 - High | 2 - Low | 2 - Medium | 2 - High |
|---|---|---|---|---|---|---|
| Enjoyment | 3.89 | 3.97 | 4.09 | 3.63 | 4.05 | 4.00 |
| Challenge | 2.89 | 2.90 | 2.86 | 2.58 | 2.40 | 2.04 |
| Interest | 3.79 | 4.02 | 3.86 | 3.60 | 3.77 | 3.72 |
| Recognition | 3.53 | 3.71 | 3.71 | | | |

| | | | | | | |
|---|---|---|---|---|---|---|
| Fairness | | | | 3.90 | 4.43 | 4.51 |
| Pass Time | 2.89 | 2.24 | 2.79 | 2.75 | 2.88 | 2.44 |

Table six: mean scores for each payment condition for six motive statements

Each survey started with five statements, with participants asked to identify the extent to which they agreed with each statement using a 5-point likert scale. Table six shows the mean score offered by participants from each experimental group for the two experiments, where 5 represents strong agreement and 1 represents strong disagreement. Feedback from participants in experiment one noted that they did not truly understand the statement “I feel recognised for my participation in the experiment” and so this was altered for experiment two and replaced with “I feel that the rewards for participating in the experiment were fair”

We compared these mean scores between the three groups for each of the two experiments, using an ANOVA test in experiment one and a Kruskal-Wallis H test in experiment two. We found **no statistically significant difference** between the groups for almost all statements. There were two exceptions to this - the pass time statement for experiment one and the fairness statement for experiment two.

For the passing time statement, the ANOVA test found a statistically significant difference between groups ($p=0.02$). The follow-up Tukey’s Honestly Significant Differences test found a statistically significant difference between the low and medium conditions only ($p=0.02$). Similarly, for the fairness statement, the Kruskal-Wallis H test found a statistically significant difference between groups ($H=13.01$, $p=0.001$). However, the follow-up Dunn’s test showed a statistically significant difference between the low and medium ($z=3.51$, $p=0.001$) and low and high conditions ($z=4.14$, $p<0.001$).

### *4.2.7 Survey Responses - Coded Motivations*

For the follow-up survey to experiment one, we asked the participants to give free text responses indicating factors which motivated their participation in the task. We then coded these free text responses and grouped them according to similar themes. Table seven shows these themes and the number of respondents who identified that motive within each of the three experimental conditions.

| **Theme** | **Low** | **Medium** | **High** |
|---|---|---|---|
| Monetary reward | 18 | 21 | 24 |
| Interest | 7 | 6 | 8 |
| Easy task | 11 | 8 | 7 |

| | | | |
|---|---|---|---|
| Desire to persevere | 6 | 11 | 3 |
| Use spare time | 6 | 2 | 2 |
| Enjoy consistency | 7 | 3 | 7 |
| Enjoy task | 1 | 8 | 7 |
| Purpose of task | 1 | 4 | 1 |
| Reminders | 5 | 4 | 1 |
| Desire to help | 0 | 0 | 3 |
| Requester Behaviour | 3 | 0 | 4 |

Table seven: Number of responses per coded theme for low, medium and high payment condition respondents to follow-up survey for experiment one

Perhaps unsurprisingly, the largest number and proportion of participants in each group identified the monetary reward as a motivation for their participation. There was a small difference in the number of participants who gave this motivation from each group, with the highest number of participants offering this motive coming from the high payment condition. However, the small number of respondents and small difference between the number of respondents means it is difficult to identify whether this difference is significant or by chance.

Beyond this, motivations generally stemmed from participants intrinsic motivations and experience of the task. A small number of participants found the task to be inherently interesting, either because of the purpose of the task, or because of the images used within the task. Participants also felt the task was relatively easy and even enjoyable -- although it should be noted that participants from the *low* payment condition appear to be significantly less likely to find the tasks enjoyable than participants from the medium and high payment conditions.

Perhaps most interestingly, however, was the impact of altruism. While response numbers were low, three participants from the high payment condition identified an altruistic desire to help with the research, a motivation which has been identified in citizen science research previously. Nevertheless, no altruistic responses were identified for the low and medium condition. Whether this is purely due to chance, the individual participants in question or a direct result of the financial incentives is unclear and given the low number of participants involved, it is not possible to explore this factor further.

As a point of comparison, for experiment two, we asked workers to rate their agreement with statements based on the coded motivations identified through experiment one. Once again, monetary rewards received the highest rating from participants, with mean scores ranging from 4.00 to 4.27. However, for this second task, participants generally disagreed that the ease of the task motivated their participation in the task, with a slightly higher rating from those in the low paid

condition, but no statistically significant difference between the three. Higher paid participants were also statistically more likely to feel an urge to persevere and complete the tasks than those in the low paid condition (p=0.03). Generally, however, participants did not strongly agree or disagree with any of the listed motivations beyond monetary rewards.

| **Theme** | **Low** | **Medium** | **High** | **KW-H p value** |
|---|---|---|---|---|
| Money | 4.00 | 4.17 | 4.27 | 0.14 |
| Enjoy | 3.24 | 3.47 | 3.50 | 0.39 |
| Difficulty | 3.10 | 2.83 | 2.81 | 0.18 |
| Persevere | 3.86 | 4.20 | 4.27 | 0.03 |
| Spare Time | 3.17 | 3.61 | 3.54 | 0.05 |
| Repeated | 3.27 | 3.78 | 3.43 | 0.09 |
| Purpose | 3.10 | 3.36 | 3.34 | 0.23 |
| Reminders | 3.44 | 4.00 | 3.84 | 0.03 |
| Requester Rep | 3.39 | 3.29 | 3.63 | 0.29 |

Table eight: Motivations, mean scores per group and Kruskal-Wallis H p-value outcomes for experiment two

# 5 Motivations for participation in the Dutch Butterfly Conservation species monitoring pilot

In this chapter, we present an analysis of one of the ACTION pilots led by DBC as an established citizen science activity. We analyse both the reasons for commencing and ceasing data collection on over 2,000 transects.

## *5.1 Methodology*

### *5.1.1 Data*

Dutch Butterfly Conservation -- an ACTION partner and pilot within WP2 -- have kept detailed records of the transects that have been set up to monitor butterfly and dragonfly species between 1990 and 2018. For each transect, there is a record of the route name and number, year in which the transect began and the reason for the transect being set up. Where applicable, the data also show the year in which monitoring of the transect ceased and the reason for this.

### *5.1.2 Method*

Using the source data, we identified a set of descriptive statistics concerning the number of active transects and total transect for each start and stop motivation. Using this data, we then calculated the correlation between individual start and stop motives. Finally, we calculated the mean number of years for which transects were monitored within the project, first accounting only for those transects for which collection has already ceased and then accounting for all transects, including active transects for which collection is still ongoing.

## *5.2 Motivations*

### *5.2.1 Commencing Participation*

| Category | Motivation | Number of transects | Number of active transects | Percentage |
|---|---|---|---|---|
| Volunteer Choice | Attractive Species | 63 | 53 | 84% |
| | Attractive Area | 470 | 200 | 43% |
| | Habitat Preference | 4 | 3 | 75% |

| | | | | |
|---|---|---|---|---|
| | Accessible Area | 196 | 87 | 44% |
| Planned Management | Species Specific | 10 | 4 | 40% |
| | Nature Management | 132 | 54 | 41% |
| | Nature Development | 86 | 42 | 49% |
| | Local Protective Measures | 42 | 15 | 36% |
| Species-Related | Species known to occur | 349 | 174 | 50% |
| | Species specific reasons | 11 | 2 | 18% |
| Request | Request of Coordinator | 345 | 183 | 53% |
| | Request of nature management organisation | 365 | 167 | 46% |
| | Other request | 160 | 54 | 34% |
| Other | Other | 133 | 95 | 71% |
| Total | N/A | 2366 | 1133 | N/A |

Table nine: breakdown of transects according to start motivation, showing total, active and proportions of transects

When analysing reasons for which transect monitoring started 2366 total transects were identified of which 733 transects resulted from intrinsic, volunteer-driven motivations and 1,633 transects resulted from extrinsic, management-driven reasons, species-specific issues or requests. Within these transects 1133 transects continued to be active at the time the data were harvested, although the proportion of transects that remain active for each individual motivation varies greatly from 18% in the case of species-specific reasons to 84% in the case of species that volunteers deemed to be interesting or attractive.

Volunteer interest in an individual area was the most significant reason for starting the observation of a transect both in terms of total and active transects, representing 20% of total transects and 18% of active transects. Conversely, habitat preference -- i.e., preference for a type of area, but not a specific location -- was the least common motivation for observations commencing. Areas in

which species were known to occur, as well as requests from coordinators or from external nature management organisations were also significant motivations leading to the commencement of transect observations.

### *5.2.2 Ceasing Participation*

| Category | Motivation | Number of participants |
|---|---|---|
| Volunteer choice | Species Decline | 25 |
| | Noise/Smell/Busy | 21 |
| | Terrain | 57 |
| | Succession/Management | 11 |
| Loss | Transect Lost | 38 |
| | Volunteer lost/moved | 80 |
| | Species disappeared | 168 |
| | Location no longer accessible | 14 |
| Availability | Volunteer has no time | 458 |
| Safety | Feeling unsafe | 5 |
| Support | Lack of support | 1 |
| Unrecorded | Unknown/volunteer death | 355 |
| Total | | 1233 |

Table ten: count of retired transects according to motivation for transect retirement

When considering the reasons for which observation of a transect ceased, there are three notable motives which led to the end of observations. The most common reason, representing 37% of all closed transects, was a change in the priorities of the volunteers or otherwise change in the amount of time that the volunteer had available. Similarly, the second largest number of transects (29%) was lost due to unspecified reasons, particularly where citizens could not be contacted to identify why they were no longer able to monitor a transect. As a result, this category is likely to overlap at least partially with a shift in volunteer priorities. Finally, the third significant category was the disappearance of the monitored species, representing 14% of transects.

These motivations are consistent across transects regardless of stop motivations. For all start motivations, volunteer availability represents the most common reason for a transect being retired,

followed by unrecorded lack of contact with volunteers. Dependent on start motivation, species disappearance, changes to terrain and volunteers moving all represent common but not frequent reasons for transect retirement.

### *5.2.2 Transect Growth and Retirement Over Time*

As figures seventeen and eighteen show, new transects have been added for monitoring each year, with a varied array of motivations driving this transect growth. Nevertheless, there have been significant changes in the nature of the motivations associated with transects at given points. There has been a significant rise in requests from coordinators and nature protection organisations in recent years, as well as in the number of transects commencing from volunteer intrinsic motivations stemming from accessible areas. Conversely, there has been a sharp fall in the number of transects started because species are known to occur. Broadly speaking, however, it should be noted that the vast majority of motivations are distributed across the lifetime of the project activities -- albeit unequally.

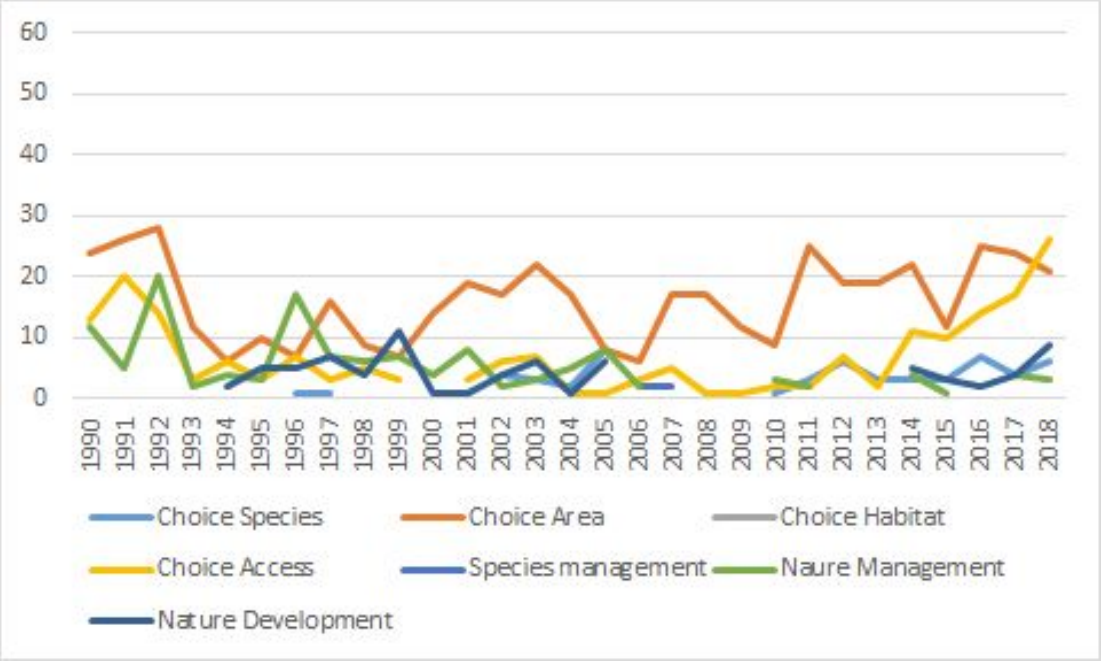


Figure seventeen: Number of transects launched each year according to start motivation (part one of two)

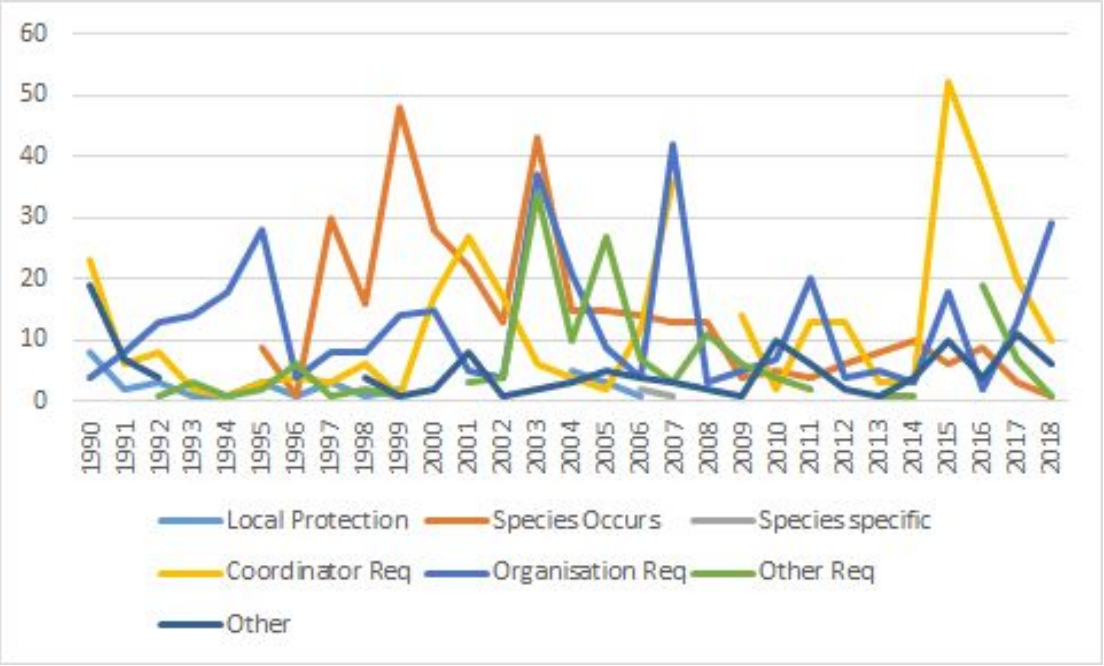


Figure eighteen: Number of transects launched each year according to start motivation (part two of two)

However, as figures nineteen and twenty show, the number of retired transects for each stop motivation is relatively low in any given year and across the lifespan of the project, with the exception of volunteer time/priorities, unrecorded loss of contact with volunteers and to a lesser extent, species disappearance.

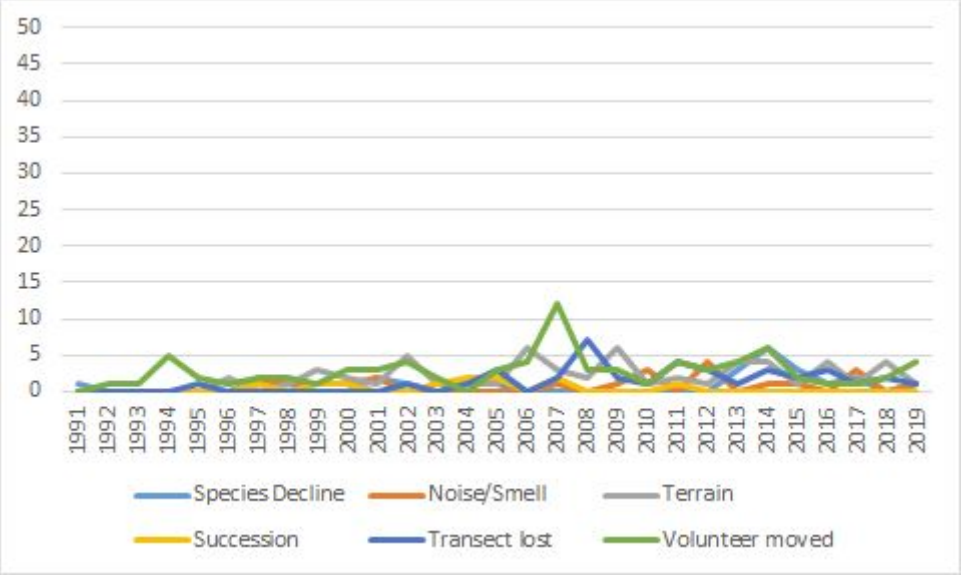


Figure nineteen: Number of transects retired each year according to stop motivation (part one of two)

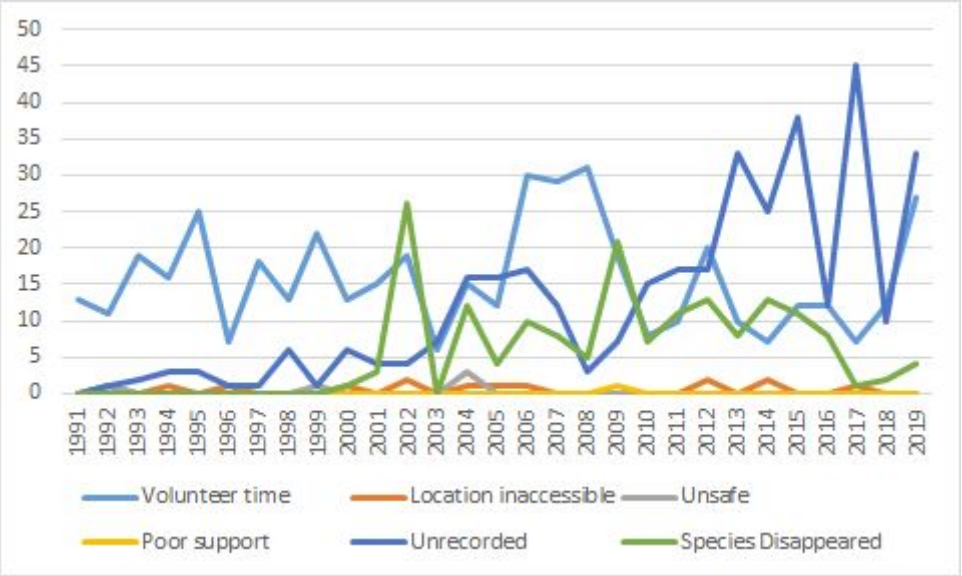


Figure twenty: Number of transects retired each year according to stop motivation (part two of two)

## *5.3 Motivations and Participation*

| Category | Motivation | Mean period (years) | Active Mean period |
|---|---|---|---|
| Volunteer Choice | Attractive Species | 1.30 | 4.49 |
| | Attractive Area | 3.60 | 4.90 |
| | Habitat Preference | 1.00 | 8.29 |
| | Accessible Area | 2.75 | 3.43 |
| Planned Management | Species Specific | 2.80 | 6.29 |
| | Nature Management | 5.99 | 9.02 |
| | Nature development | 4.34 | 9.10 |
| | Local Protective Measures | 8.86 | 10.57 |
| Species-Related | Species Known to Occur | 5.69 | 7.40 |
| | Species Specific Reasons | 4.55 | 5.92 |
| Request | Request of Coordinator | 2.48 | 4.69 |
| | Request of nature | 3.53 | 6.02 |

| | | | |
|---|---|---|---|
| | management organisation | | |
| | Other Request | 4.05 | 4.95 |
| Other | Other | 2.17 | 7.03 |

Table ten: mean number of years transects were monitored grouped by start motivation. Note that Active Mean Period includes all active transects, based on the number of years they have been gathered and may not be an underestimate of how long a transect will be gathered for.

Table ten shows the length of time for which transects were completed, grouped according to the initial motivation that led to the commencement of the transect completion. We include the mean number of years for which a transect was collected including only closed transects and all transects, both active and inactive.

Findings suggest that generally intrinsic, volunteer-driven reasons for participation result in less strong participation motivations and therefore, briefer periods of participation than extrinsic, management and species-related reasons. While the habitat preference category is associated with a very high mean contribution time of 8.29 years, this is largely a result of the low number of participants involved and the high number of participants who continue to observe transects within this category. When accounting only for transects which are no longer observed, this category had the lowest mean contribution time, representing just one year.

In contrast, extrinsic motivations are associated with higher mean participation periods, whether accounting for all transects or only inactive transects. In particular, planned management activities -- in particular nature management and nature development -- are associated with the highest mean contribution times, representing 9.02 and 9.10 years on average. Similarly, local protective measures feature the highest participation periods, whether considering only retired transects or all active transects, with 8.86 and 10.57 years respectively. On the hand, it should be noted that this may partially result from the time taken to complete these activities. Management activities and other extrinsic factors including requests may be more essential and necessitate long-term observations, regardless of whether other issues arise -- for example, species disappear or minor discomfort issues which might lead a volunteer to stop observing other transects.

There are, however, two significant limitations with these data. The first is that many transects are still active and there is no consistent date on which individual transects were launched. As a result, it is difficult to accurately compare active transects, which vary in start year from 1990 to 2018. Similarly, it is important to note that the number of transects from which figures were calculated is in many cases very low.

# 6 Guidelines

From our findings, we summarise the following conclusions:

- Citizen Scientists tend to be more invested in tasks (and by extension, associated research) than other crowdsourcing participants
- Citizen Scientists tend to have stronger and more diverse motives than other crowdsourcing participants
- Citizen scientists tend to be encouraged to initially participate in citizen science activities due to their own *intrinsic* motivations
- Nevertheless, these intrinsic motivations are not necessarily associated with long term engagement
- Paid crowdsourcing can result in valuable scientific research data
- Paid participants can continue to participate in the medium- to longer-term if paid and rewarded accordingly
- Higher levels of pay do not result in *better* or even greater volumes of submissions
- Higher levels of pay are associated with increased likelihoods of poor quality, low-effort submissions
- Other extrinsic factors -- e.g., the needs of scientists or other stakeholders and organisational requirements -- can also motivate consistent participation from participants, even if not derived from the participants' own interests

On the basis of these findings, we derive a set of preliminary guidelines for citizen science projects:

- Emphasise intrinsic motivations for new participants, to maximise initial participation and interest within projects. This may be achieved through the framing of tasks, advertisements or task documentation or through communication with participants.
- Introduce extrinsic motivations derived from project/researcher needs and expectations at an early stage. This may be through feedback indicating how important participant contributions are, or updates on how the project is proceeding, with an emphasis on the effort required by the project. This will help to encourage longer-term participation deriving from altruism as intrinsic motivations become less effective.
- If necessary, paid participation can boost data gathering processes, where participation may otherwise be low. Carefully consider the differing motivations of the two participant profiles however -- paid participants have less interest in the research, task or its purpose and are primarily or solely motivated by payment. At the same time, providing too large a payment payment risks motivating undesirable behaviours and poor quality data.
- Wherever possible, design projects to appeal to a diverse range of motivations and interests. Citizen science participants are not homogeneous and while participant motivations can broadly be broken down into intrinsic and extrinsic motivations, individual motivations and the effectiveness of these motivations within these groupings vary significantly.

# 7 Future Work

In this section, we outline ongoing and future research by the ACTION team in the context of incentives and motivations. This research will be presented in D5.7, the final analysis and guidelines for incentives and motivation in citizen science and will be used to formulate the final guidelines for projects.

## *7.1 Understanding motivations in ACTION pilots*

We will work further with ACTION pilots to understand motivations in pollution-specific activities and identify how -- if at all -- these differ from motivations in other forms of citizen science. Accelerator pilot CitiComPlastic has developed a reflection methodology to gather data around understanding *what*, *why* and *how* their participation has impacted them. This leads to the following set of questions:

- What did you do and Why? What did you expect? What went well?
- What impact did the experience have on you? Why does it matter? How did this experience clarify, expand, or create an interest?
- How will you apply what you have learned? How does this move you towards your vision of the future?

Notes from these sessions will be shared with ACTION and we will use these to develop a greater understanding of motivations, as well as methodologies to gather data from other ACTION pilots.

## *7.2 Incentives for pollution crowdsourcing*

KCL have developed a platform to deliver specialised crowdsourcing activities embedded within a completely customisable interface known as QrowdSmith. This platform offers two major affordances for ACTION: the first is to deliver more customisable tasks within an easily editable framework. This framework allows us to alter the way that tasks are framed to focus more or less on scientific issues, pollution and specific motivations such as altruism, intrinsic motives and extrinsic or game-like features.

The second is to completely customise the incentives offered to participants. This could be as simple as offering feedback or as complex as introducing a gamified leaderboard and point system to track participation. Qrowdsmith features a removable chat plug-in to allow more complex *social* incentives and interaction, which have been suggested to impact participation in previous citizen science studies (Reeves et al, 2017; Reeves and Simperl, 2019).

Qrowdsmith is already complete and under evaluation. We hope to make use of the platform to develop citizen science and pollution-specific tasks, which will be compared with more general crowdsourcing activities to understand motivations specific to these contexts and how they differ from motivation for participation in crowdsourcing more generally.

# 8 Conclusions

In this deliverable, we have outlined research into the motivations associated with volunteer participation in citizen science in the domain of pollution and the effectiveness of financial incentives in motivating participation. We carried out a survey to identify factors that motivated participation within a photometer network for monitoring light pollution and compared these with participation in crowdsourcing more broadly. Additionally, we analysed motivations for commencing and ceasing participation in a butterfly counting citizen science project which monitors the impacts of chemical pollution such as pesticide on different species. In terms of incentives, we explored the impact that financial incentives have on a light pollution monitoring project, to understand how continued participation is impacted by these payments.

In future work, we will explore motivations in further pollution based citizen science projects drawn from ACTION pilot projects. We will further conduct additional experiments to monitor the effectiveness of distinct incentives such as gamified rewards and social incentives and pressures within pollution monitoring citizen science.

# 8 References

Aoki, P., Woodruff, A., Yellapragada, B., & Willett, W. (2017, May). Environmental protection and agency: Motivations, capacity, and goals in participatory sensing. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (pp. 3138-3150).

Berg, J., Furrer, M., Harmon, E., Rani, U., & Silberman, M. S. (2018). Digital labour platforms and the future of work. *Towards Decent Work in the Online World. Rapport de l'OIT*.

Buytaert, W., Zulkafli, Z., Grainger, S., Acosta, L., Alemie, T. C., Bastiaensen, J., ... & Foggin, M. (2014). Citizen science in hydrology and water resources: opportunities for knowledge generation, ecosystem service management, and sustainable development. *Frontiers in Earth Science*, *2*, 26.

Commodore, A., Wilson, S., Muhammad, O., Svendsen, E., & Pearce, J. (2017). Community-based participatory research for the study of air pollution: a review of motivations, approaches, and outcomes. *Environmental monitoring and assessment*, *189*(8), 378.

Hara, K., Adams, A., Milland, K., Savage, S., Callison-Burch, C., & Bigham, J. P. (2018, April). A data-driven analysis of workers' earnings on Amazon Mechanical Turk. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (pp. 1-14).

Kinchy, A. (2017). Citizen science and democracy: Participatory water monitoring in the Marcellus shale fracking boom. *Science as Culture*, *26*(1), 88-110.

Kraut, R. E., & Resnick, P. (2012). *Building successful online communities: Evidence-based social design*. Mit Press.

Lafourcade, M., Joubert, A., & Le Brun, N. (2015). *Games with a Purpose (GWAPS)*. John Wiley & Sons.

Levontin L., C., Gilad: Questionnaire for the Motivation for Citizen Science Scale. Technical report, Technion, Israel Institute of Technology (2018)

Lindemann, N.: What's the average survey response rate? [2019 Benchmark]. Technical report, Survey Anyplace (2019), https://surveyanyplace.com/average-survey-response-rate/

Maisonneuve, N., Stevens, M., & Ochab, B. (2010). Participatory noise pollution monitoring using mobile phones. *Information polity*, *15*(1, 2), 51-71.

Mao, A., Kamar, E., Chen, Y., Horvitz, E., Schwamb, M. E., Lintott, C. J., & Smith, A. M. (2013, November). Volunteering versus work for pay: Incentives and tradeoffs in crowdsourcing. In *First AAAI conference on human computation and crowdsourcing*.

Ponciano, L., Brasileiro, F., Simpson, R., & Smith, A. (2014). Volunteers' engagement in human computation for astronomy projects. *Computing in Science & Engineering*, *16*(6), 52-59.

Reeves, N., Tinati, R., Zerr, S., Van Kleek, M. G., & Simperl, E. (2017, February). From crowd to community: a survey of online community features in citizen science projects. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing* (pp. 2137-2152).

Reeves, N. T., & Simperl, E. (2019). Efficient, but Effective? Volunteer Engagement in Short-term Virtual Citizen Science Projects. *Proceedings of the ACM on Human-Computer Interaction*, *3*(CSCW), 1-35.

Reeves, N., West, P., & Simperl, E. (2018, June). A game without competition is hardly a game": The impact of competitions on player activity in a human computation game. In *Sixth AAAI Conference on Human Computation and Crowdsourcing*.

Richter, A., Hauck, J., Feldmann, R., K ¨uhn, E., Harpke, A., Hirneisen, N., Mahla,A., Settele, J., Bonn, A.: The social fabric of citizen science—drivers for long-term engagement in the German butterfly monitoring scheme. Journal of insect conservation 22(5-6), 731–743 (2018) https://doi.org/10.1007/s10841-018-0097-1

Rotman, D., Preece, J., Hammock, J., Procita, K., Hansen, D., Parr, C., ... & 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).

Scandolari, D., Re Calegari, G., Scrocca, M., Celino, I.: TESS Network Motivation Survey Dataset (May 2020). https://doi.org/10.5281/zenodo.3739058

Schwartz, S.H.: An overview of the schwartz theory of basic values. Online readings in Psychology and Culture 2(1), 2307–0919 (2012)

Strasser, B., Baudry, J., Mahr, D., Sanchez, G., & Tancoigne, E. (2019). " Citizen Science"? Rethinking Science and Public Participation. *Science & Technology Studies*, *32*(ARTICLE), 52-76.

Tweddle, J. C., Robinson, L. D., Pocock, M. J. O., & Roy, H. E. (2012). *Guide to citizen science: developing, implementing and evaluating citizen science to study biodiversity and the environment in the UK*. NERC/Centre for Ecology & Hydrology.

Woodcock, J. C., Greenhill, A., Holmes, K., Graham, G., Cox, J., Oh, E. Y., & Masters, K. (2017). Crowdsourcing citizen science: Exploring the tensions between paid professionals and users. *Journal of Peer Production*, (10).