University of Southampton
Faculty of Physical Sciences and Engineering
Web Science

Sociality, Community and Productivity in Virtual Citizen Science

Neal Reeves

Thesis for the Degree of Doctor of Philosophy

August 2018

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF PHYSICAL SCIENCES AND ENGINEERING

Thesis for the degree of Doctor of Philosophy

SOCIALITY, COMMUNITY AND PRODUCTIVITY IN VIRTUAL CITIZEN SCIENCE

Neal Reeves

Virtual Citizen Science describes web-based crowdsourcing activities which recruit volunteers to complete microtasks for scientific research. VCS methodologies have been applied to diverse research challenges, from identifying neurons in MRI-scan images of the optic nerve, to morphologically classifying images of galaxies. Initiatives generally rely on participants' intrinsic motivations to encourage contributions, but increasingly designers are turning to additional mechanisms – including the use of online community features and discussion platforms such as forums. However, the influence of these features on participant engagement are still poorly understood. Drawing on a pragmatist, mixed-methods approach, this thesis explores the relationship between these social features within task and discussion elements of projects, to understand the influence of such features on volunteer productivity and project efficiency. A literature review of five transdisciplinary databases was conducted to identify design principles, motivations and social features associated with VCS approaches. This was then followed with a review of 48 VCS projects, to better understand the online community features within current VCS initiatives. To understand and clarify these findings, interviews were conducted with six members of the EyeWire project design team. Analysis of competitions in EyeWire demonstrated a relationship between task sociality and increased productivity and activity within the project. Finally, an analysis of two high-pressure VCS projects explores how periods of heavy productivity affect discussion activity. This research contributes to the understanding of motivational factors and design affordances within Virtual Citizen Science and similar crowdsourcing initiatives.

Contents

List of Tables

List of Figures

Declaration of Authorship

I, Neal Reeves, declare that this thesis, entitled "Sociality, Community and Productivity in Virtual Citizen Science", and the work presented in it are my own and has been generated by me as the result of my own original research. I confirm that:

1. This work was done wholly while in candidature for a research degree at this University;
2. Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
3. Where I have consulted the published work of others, this is always clearly attributed;
4. Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
5. I have acknowledged all main sources of help;
6. Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
7. Parts of this work have been published or otherwise disseminated as:
    (a) Manuscripts:

        Neal Reeves, Ramine Tinati, Sergej Zerr, Elena Simperl, and Max Van Kleek. 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, pages 2137–2152. ACM, ACM, 2017

Neal Reeves, Peter West, and Elena Simperl. ‘`a game without competition is hardly a game”: The impact of competitions on player activity in a human computation game. In *Proceedings of the AAAI Conference on Human Computation (HCOMP 2018)., HCOMP*, volume 18, 2018

Elena Simperl, Neal Reeves, Christopher Phethean, Todd Lynes, and Ramine Tinati. Is virtual citizen science a game? *Transactions on Social Computing*, 2018

(b) Talks:

Neal Reeves, Max Van Kleek, and Elena Simperl. From crowd to community: Support for community features in online citizen science projects. Presented at GESIS Computational Social Science Winter Symposium 2015 in Cologne, Germany on 2 December 2015, 2015

Neal Reeves. Sociality in virtual citizen science (presentation and panel discussion). Presented at 20th ACM Conference on Computer-Supported Cooperative Work and Social Computing in Portland, OR, USA on 25 February 2017, 2017

Neal Reeves, Ramine Tinati, Sergej Zerr, Elena Simperl, and Max Van Kleek. 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*, pages 2137–2152. ACM, ACM, 2017

Signed: ....

Date: 2nd August 2019

Acknowledgements

I would like to thank my supervisors, Elena Simperl and Jeff Vass for their support, guidance, encouragement and comments throughout my research. I am very grateful for the valuable advice and support that has been offered, both in terms of the research itself, but also in terms of events and conferences to attend, opportunities to apply for, publications to submit to and how best to navigate the PhD and associated academic processes. I certainly would not have got here without it.

I'd also like to thank everybody at the Web Science DTC/CDT who gave me this opportunity and made my time so enjoyable, both for the administrative support but also for the barbeques, parties, meet-ups and discussions. I'm especially grateful for being given the funding and necessary knowledge to pursue this research and – at times – for the various pressures placed on me to make sure I finished it.

As well as the administrative staff and academics, I'd like to thank my fellow PhD students, particularly within my cohort, for all of the support, feedback and assistance with questions when things were going wrong. It made going through this much less painful and without you all I might still be wrestling with a broken 9-month report template.

I would like to thank everyone at the EyeWire team not only for taking part in the interviews and offering data for my research, but also for being such gracious hosts in Boston, even in the middle of February in the midst of a snowstorm. I'm particularly grateful for their patience as 20-minute interview sessions turned into 2 hour marathons and for their stamina in still being willing to go out for pizza at the end. Thank you also to Grant from the Zooniverse for helping me through the Stargazing Live data and answering numerous questions despite quite literally being on an arctic expedition.

To my long-suffering friends, I'd just like to say I appreciate you putting up with me for the last four years – suffering through me discussing my PhD, discussing everything that's going wrong with my PhD and attempting to do – and discuss – anything other than my PhD. Hopefully this will be the end of it.

And finally thank you Pete, for everything over the course of this PhD and before.

Chapter 1

Introduction

Modern technologies pose significant challenges to traditional scientific workflows. In an age characterised as the "age of big data", businesses and governmental organisations generate and store far larger volumes of data than ever before (Tene and Polonetsky, 2012). This is equally true of the tools and technologies used in modern scientific endeavours: from the Sloan Digital Sky Survey which produces hundreds of thousands of images annually of astrophysical phenomena in space, to the Large Hadron Collider which generates 60 terabytes of data every day (Bryant et al., 2008; Lintott et al., 2008). Existing tools and methods for analysing such large volumes of data for scientific research purposes are often insufficient, being too inaccurate and requiring more time and resources than are available to the small teams of scientists traditionally responsible for such analysis processes (Lintott et al., 2008).

One proposed solution to this issue is *Citizen Science* (CS), which comprises the recruitment of volunteer 'citizen scientists' to assist with gathering or analysing data in scientific research. The involvement of volunteer non-professionals in science is not a recent phenomenon; until the nineteenth century, the majority of research was conducted by non-professionals (Miller-Rushing et al., 2012). These early citizen scientists collected and analysed scientific data as a hobby, while pursuing other forms of employment (Silvertown, 2009). Over time, and as research became the prevail of trained experts, opportunities for volunteer involvement in scientific research persisted, predominantly through volunteers interacting with scientists in small-scale, localised projects (Catlin-Groves, 2012). With the advent of the World Wide Web and the uptake of approaches such as open science and crowdsourcing, projects have increasingly allowed volunteers to use online tools to engage in professional research on a much larger scale (Bonney et al., 2009; Kobori et al., 2016). These *Virtual Citizen Science* projects – CS activities conducted over the web (Wiggins and Crowston, 2011) – commonly divide large scientific research goals into smaller individual *microtasks* to be completed amongst communities numbering tens or hundreds of thousands of volunteers (Tinati et al., 2015b). Each volunteer can contribute a small part of the scientific research puzzle by completing individual and repeatable microtasks (Tinati et al., 2015b). Indeed, these projects have often been described as being synonymous with 'crowdsourced science' (Johnston et al., 2017; Wiggins and Crowston, 2011).

To support volunteers in completing microtasks, projects are increasingly implementing discussion platforms such as message board forums to facilitate task completion and provide opportunities for interaction between volunteers, project scientists and administrators (Jennett et al., 2013). However, sources in the wider literature are contradictory with regard to the nature of Citizen Science as a social, community-based activity. Johnston et al. (2017) describe amateur participation in science as having been "liberated by electronic media", both in terms of opportunities for access to scientific activity and for discussion and organisation around these topics. Nonetheless, the authors also describe CS projects as strongly scientist-led, with restrictions which prevent volunteers from accessing data and interacting with one another, requiring the use of external social media platforms (Johnston et al., 2017). Conversely, while Wiggins and Crowston (2011) agree that projects are scientist-led, their definition of Citizen Science details the common use of social elements such as competition and collaboration.

Other definitions of CS indicate a more diverse field. In Haklay's typology, projects range between highly restrictive volunteer computing projects, where participants offer computing power or sensor data, to 'extreme' Citizen Science, where participants are a fundamental part of the entire scientific process (Haklay, 2013). The importance assigned to community interaction varies between platforms and projects; Tinati et al. (2015b) summarise the views of the Zooniverse design team, who prioritise interaction and communication as essential parts of Citizen Science workflows. Yet, when also discussing Zooniverse, Mugar et al. (2014) describe that the scientific value which may arise from *limiting* community interaction includes ensuring the potential accuracy of submissions.

There has been a rapid expansion in the number and variety of Citizen Science projects in recent years, fuelled in part by the adoption of new technologies by scientists and volunteers alike (Catlin-Groves, 2012; Kullenberg and Kasperowski, 2016). Virtual Citizen Science (VCS) is no exception, with efforts such as the Zooniverse's *Panoptes* service now allowing anyone to launch a VCS project (Bowyer et al., 2015; Kosmala et al., 2016). Moreover, sources of funding for scientific research have been stretched increasingly thin and Citizen Science has been proposed by some as a potential solution to this due to its efficiency (Sparks et al., 2015). However, the funds made available for CS research are not always compatible with the variety of skills required to effectively run a project for any significant period of time, nor the range of additional activities required of CS initiatives, such as the need for outreach and dissemination (Haklay, 2015; Freitag and Pfeffer, 2013). In some cases, projects may have no funding at all (Catlin-Groves, 2012). This in turn puts pressure on VCS projects to boost and maximise productivity to make the best use of the limited funds available (Celino et al., 2018). Yet the push for efficiency requires an understanding of the complex motivational factors and influence of various affordances on volunteer activity (Lintott and Reed, 2013). It is perhaps unsurprising, then, that improving the efficiency of VCS projects is a key area of current research (Tinati et al., 2015a).

In light of these concerns, the decision to introduce and promote the use of discussion and community activities in VCS projects may seem counterintuitive; these activities lead to volunteers spending valuable time engaging in activities other than classification and contribution (Tinati et al., 2015a). Moreover, community activities must be moderated and time must be spent responding to questions and queries; this was an issue which the introduction of discussion forums to Galaxy Zoo was in fact originally intended to circumvent (Tinati et al., 2015b). But is it correct to think of a task and discussion dichotomy? Is there a relationship between discussion activity and task completion, productivity and overall efficiency? If the view of VCS presented by Mugar et al. (2014) is accurate, should VCS even be considered a social activity? This thesis will approach these issues with an interdisciplinary, mixed-methods Web Science approach. Drawing on a transdisciplinary review of contemporary peer-reviewed literature, interviews with project designers and administrators and analysis of data from VCS projects, I will explore the relationship between community interaction and VCS contributions.

1.1 Research Questions

The aim of this thesis is to explore the nature of sociality and social interaction in Virtual Citizen Science and how social activity relates to task productivity, as well as to consider VCS projects as examples of online communities. The overarching research question of this thesis is therefore as follows:

> *What is the relationship between task activity and social activity within Virtual Citizen Science?*

However, such a question alone would be insufficient to accurately and fully explore any potential relationship between social and task activity. The form that social features take in VCS is unclear and, as demonstrated above, there is some ambiguity in the classifications of VCS as a social- or individual-driven process. Moreover, with the rapid expansion of Citizen Science generally and VCS specifically in recent years, it is possible and even probable that the core characteristics and features of VCS have shifted. As a result, the overall research question has been broken down into three smaller research questions, one of which has been broken down further into two complimentary sub-questions:

1. *How is Virtual Citizen Science understood and realised today?*
   Before discussing any relationship within Virtual Citizen Science, it is first essential to outline what VCS actually entails, particularly in the light of the recent growth within the field and the adoption of new tools and technologies. This question aims to understand the characteristics and processes involved in VCS projects, the associated design features and challenges within VCS initiatives and the characteristics of volunteer engagement. The results and outcomes of this question will inform the methods and data used in subsequent questions, such that the focus of this thesis is to be on *Virtual* Citizen Science in its current form, rather than considering outdated or irrelevant projects and concepts.

2. *What online community tools and features are present in Virtual Citizen Science projects?*
   Given the ambiguity and contradictions identified above in terms of the social features and role of the community, it is necessary to supplement the understanding of VCS achieved through question 1 with an understanding of the associated tools and features and how these allow volunteers to view

and interact with the wider community. This question seeks to outline the online community features that are involved in VCS projects, particularly within task elements and interfaces. What features are available to participants and what role do these play within projects?

3. *How do task-based social events influence task completion in VCS projects?*
   Virtual Citizen Science activities can broadly be understood in terms of two distinct, yet interrelated element. The first of these two elements concerns the crowdsourced microtask assigned to participants and by extension the features associated with that element – for example, gamified elements, tutorials, rewards and feedback. This question seeks to understand what – if anything – are the effects of social events such as competitions attached to this task-based element on the productivity of individuals and overall task completion in VCS projects.

4. *To what extent is engagement in discussion associated with increased engagement in task completion?* If task-based activities and features represent one core element of VCS projects, then discussion and social features represent the other. The aim of this question is volunteers who make the (purely voluntary) decision to engage with these discussion opportunities tend to complete more tasks than those volunteers who don't. The formulation of this question is informed in part by the the overarching research question, but also from the findings of questions 1-4.

1.2 Research Contributions

In researching the research questions stated in the previous section, this thesis makes five core contributions:

1. *Carry out and present the findings of a large-scale literature review centred on Virtual Citizen Science, drawn from a transdisciplinary set of sources.* Chapter 4 presents findings from a literature review drawn from five databases. I identify a framework of motivational factors and affordances associated with volunteer participation in VCS activities, with a focus on both talk and task elements.

2. *Develop a theoretical framework of four themes and guidelines for online community design.*
   Based on online communities literature, in chapter 5 I develop an evidence-based framework of themes and guidelines associated with productivity in online communities, which I use to explore and evaluate the design features present in a number of VCS projects.

3. *Identify the role played by sociality in influencing the design of a VCS project.*
   Drawing on phenomenological interviews with six members of the EyeWire CS project team, chapter 6 explores the factors leading to the decision to introduce sociality and social features, both in terms of the EyeWire task and in terms of the live-chat, forum and Wiki services associated with the project.

4. *Identify the impact of adding sociality to a Virtual Citizen Science task.*
   Using evidence from the EyeWire project and supported by findings from chapter 6, chapter 7 explores the impact of adding social competition activities to VCS tasks, using hypothesis testing and effect sizes to identify what effect the addition of sociality has on productivity and participant numbers.

5. *Identify the interaction between talk and task in two high-pressure VCS projects.*
   Chapter 8 explores the relationship between talk and task activity over time and across participants in two high-pressure projects designed to run for approximately 48 hours (a relatively new and uncommon format for a VCS project to take).

1.3 Research Scope and Generalisability of Findings

Citizen Science is a diverse and growing field and this is no less true for Virtual Citizen Science as a subset of this field. The aim of this thesis is to – where possible – consider the impact of sociality and social features on productivity across VCS projects and within VCS approaches. Nevertheless, it is important to define the scope of this research. There is no single repository or source from which individual VCS projects can be identified and there remain areas of disagreement as to the characteristics and nature of VCS as distinct from other, similar citizen science initiatives. For example, there remains debate about whether volunteer computing should be considered an example of Citizen Science activity; compare, for example, Haklay (2013) and Wiggins and Crowston (2012). Similarly, Kullenberg and Kasperowski (2016) identified 490 citizen science projects through the Web of Science platform and yet few, if any, of these projects correspond to the criteria that generally define a VCS project (see chapter 4 for more details.)

Given these factors, it would be infeasible as part of this thesis to consider every possible Virtual Citizen Science project. Moreover, any such analysis may well quickly become outdated as new projects are launched and old projects are retired. Instead, I will ensure generalisability through two strategic processes. The first of these is to use a more large-scale overview of the VCS field through the literature and project review processes discussed in chapters 4 and 5. This process involves accounting for a diverse variety of projects to identify commonalities and points of departure in the implementation and understanding of VCS and to define common principles which inform the findings of subsequent chapters. This large-scale overview, then, accounts for the state of the VCS field at the time this thesis was written and informs the selection of projects to be selected for further analysis later. In turn, the findings of the literature and project review process allow for the identification of features and characteristics that are typical of VCS projects and generalisability will thus further be enforced by selecting projects that adhere to these findings and which accurately represent the wider VCS landscape for use throughout the interview and data analysis processes. Specific findings relating to these issues are discussed in more detail in chapters 2 and 3 and a further clarification regarding the scope of this thesis is offered in chapter 4. Nevertheless, a typical VCS project can be briefly summarised as one that relies predominantly on altruistic motivations and volunteer effort from those with high levels of intrinsic motivations, with the potential for virtual – but generally not tangible – rewards, which assigns microtask crowdsourcing activities to volunteers for completion and which otherwise offers low levels of autonomy. VCS projects largely belong to one of two main groups – gamified projects or 'Games With A Purpose', which use game elements or game-like activities to encourage task completion and non-gamified projects in which the task itself is its own reward, with few extrinsically motivating features. Throughout this thesis, I will – wherever possible – consider the differences between these project types and their associated features and design decisions.

1.4 Structure

The remainder of this thesis is set out as follows:

- Chapter 2 presents Citizen Science in more detail, identifying key typologies and categories that have been applied to this field and conceptualises and differentiates VCS from other related initiatives. In addition, theories and concepts which underpin the methodological processes and analyses conducted throughout this thesis are presented in detail, with reference to related work from the literature where necessary.

- Chapter 3 details the methodological approach followed in chapters 4 to 8. The philosophical paradigm of *pragmatism* is discussed and the *mixed methods* approach is explained, as well as the data collection processes followed throughout each additional chapter. Moreover, the specific qualitative and quantitative methods and methodological concepts for this research are selected, discussed and justified in the context of the characteristics of the collected data.

- Chapter 4 presents a systematic review of VCS literature across three themes: design challenges and solutions, participant motivation and engagement, and sociality and interaction.

- Drawing on the results of the systematic literature review and additional online communities literature, chapter 5 explores the inclusion of online community features in VCS platforms. Drawing on a survey of 48 VCS projects, I explore four themes identified from the literature as key to motivating participation: task visibility, goals and challenges, feedback and rewards.

- To contextualise and further explore the findings of the project review, chapter 6 presents the results of phenomenological interviews with six members of the EyeWire CS project team. These interviews explore the reasons for introducing specific social features to VCS, as well as present anecdotal evidence and observations regarding the effectiveness of particular project characteristics.

- Chapter 7 presents statistical analysis and hypothesis tests exploring and quantifying the effectiveness of competitions which introduce a level of sociality into the EyeWire project task.
- Chapter 8 presents an analysis of the relationship between task completion and sociality within two short-term, high pressure projects, conducted over the course of five days and three days respectively.
- Finally, chapter 9 synthesises and summarise the findings, arguments and issues raised within the previous chapters, drawing conclusions and identifying limitations and areas for future research.

Chapter 2

Background

This chapter describes several important concepts for this research, beginning with an outline of Citizen Science as a field and highlighting the characteristics that distinguish *Virtual* Citizen Science from other, more traditional forms. This is followed by an explanation of central concepts which will be discussed or considered elsewhere in this thesis, namely *online communities*, *gamification* and *crowdsourcing*. Finally, three central sociological and psychological concepts with importance for human computation and human computer interaction are presented, specifically *collective action*, *distributed cognition* and *social machines*.

2.1 Citizen Science

Citizen Science (CS) describes the engagement of volunteers in professional research, either in collaboration with or on behalf of professional researchers (Bonney et al., 2009). The term *volunteer* here refers to any individual who engages with the project and contributes their effort without being paid. Generally speaking, there is a distinction between project scientists (who may also be unpaid volunteers, particularly in the case of PhD students or interns) and the project participants who are volunteers or *citizens* (Eitzel et al., 2017). Nonetheless, *volunteers* correspond to different stakeholder groups – from indigenous or local groups to uninvested observers – and have different levels of involvement, contributing money, materials, data or time (Eitzel et al., 2017). These volunteers also possess variable levels of knowledge about the scientific domain to which they contribute, but a common feature of CS is that projects leverage – and participants contribute – participant expertise (Eitzel et al., 2017;

Irwin, 2015). In the simplest sense, the term *volunteers* can here be understood to describe anyone who is "not affiliated with credentialed academic/research institutions" (Kimura and Kinchy, 2016).

In spite of the term Citizen *Science*, such activities do not necessarily describe or require the completion of scientific research as a primary goal (Silvertown, 2009). Citizen Science projects have a strong potential for educating participants in scientific techniques and concepts and are therefore used in both field and classroom settings (Kloetzer et al., 2014; Kridelbaugh, 2016). Moreover, CS is a highly diverse field, with distinct sub-types that – while arguably of benefit to scientific research – do not necessarily entail the gathering of data for publication or associated research applications. While CS is not the only form of volunteer involvement in scientific activities, it is distinct in that the number of participants, the geographic spread of data sources and the quantity of data involved are all significantly higher within CS (Wynn, 2017). This is particularly the case in technology-enabled forms of CS, which are "unique," due to their exploitation of "widespread computing resources" and the Web "which connects them" (Wynn, 2017).

In fact, research-oriented CS activities describe a minority of the CS field as a whole. In a meta-analysis completed by Kullenberg and Kasperowski (2016), which analysed 2,568 CS-related publications and identified 490 CS projects drawn from the Web of Science database, just 16% of projects stated specific scientific research goals. A likely contributing factor to this disparity is the slow rate of data publication associated with CS project. Formal publication of data and findings often occurs long after the completion of the data gathering or analysis process and the retirement of projects, if indeed such data are published at all (Theobald et al., 2015).

To illustrate and classify the diverse forms of CS, Wiggins and Crowston (2011) divide CS projects and initiatives into a framework of five types, distinguished through a combination of factors including task type, location, project aim and the nature of the roles of project administrators and volunteers. In actuality, only three of these sub-types have explicit scientific applications, of which only two have scientific research and publication as their end goal:

- **Action**: Citizen-led interventions based on local environmental issues. Overall aim concerns a localised, tangible community issue such as monitoring and improving local water quality.

- **Conservation**: Scientist-led projects aiming to encourage awareness and stewardship of natural resources.
- **Education**: Scientist-led or educator-led projects where education and outreach are the primary, rather than secondary, goals.
- **Investigation**: Scientist-led projects requiring data collection from a physical environment, at a local, national or international level.
- **Virtual**: Scientist-led projects where all project activities are entirely conducted virtually, independent of the physical environment.

Where projects *do* involve engagement in scientific research, the nature and extent of this engagement is also highly variable. Although the specific tasks and activities involved in any given project are highly context- and project-dependent, initiatives can be broadly categorised according to the stages of the scientific research process in which volunteers engage and the level of freedom they have to do so (Haklay, 2013; Wiggins and Crowston, 2012):

1. **Citizen Sensing and Volunteered Computing**: Participant engagement is momentary and autonomy is low. Participants install a sensor or software on a smartphone or computer, through which data is gathered and have little engagement with project activities beyond this.
2. **Distributed Intelligence/Contributory**: Participants carry out simple interpreting, analytical or data gathering tasks. May involve dissemination of findings.
3. **Participatory Science**: Participants contribute to the definition of the research problem or question and carry out more complex data collection and analysis tasks than in citizen sensing/distributed intelligence initiatives.
4. **Collaborative**: Volunteers engage in data collection, analysis and potentially study design, interpretation, draw conclusions and disseminate results.
5. **Extreme Citizen Science/Co-created Citizen Science**: Volunteers engage in all stages of the scientific process. Unlike other stages, which are scientist-led, extreme Citizen Science projects either entail collaboration

between scientists and volunteers or may be led by individual volunteers or communal efforts.

This typology assumes that participant engagement does not expand – that is, a contributory or distributed intelligence project may allow participants to disseminate findings, but there is no realistic opportunity or possibility that engagement will expand to other stages. Extreme or Co-Created Citizen Science projects are rare at the time of writing this thesis; when conducting a literature review on the topic of Virtual Citizen Science, no evidence was found of any project which could be described in such terms[1] and neither Haklay (2013) nor Wiggins and Crowston (2012) offer examples of specific projects that may correspond to these categories. In contrast, while citizen sensing and volunteered computing projects are common (Catlin-Groves, 2012), there is some debate about whether these should be considered CS initiatives. Notably, Wiggins and Crowston (2012) do not include such initiatives in their typology, which defines contributory CS as active engagement in the data gathering or analytical process, rather than a passive process driven by the installation of software. A similar distinction is made by Eitzel et al. (2017), who discuss an active and conscious process, where volunteers actively and knowingly contribute data or resources to a scientific endeavour, for scientific research purposes.

A more cross-cutting, if simple, definition framework is described by Tinati et al. (2015b), who divide projects into one of three high level categories based on the main activities in which participants engage: *data collection*, *data analysis* and *problem solving*. While problem solving is implied to be a separate process or task type, it also underpins data collection and data analysis initiatives[2]. Additionally, the terms *data collection* and *data analysis* have a different meanings to those of the typologies of Haklay (2013) and Wiggins and Crowston (2011, 2012). Tinati et al. use *data collection* to explicitly refer to the active collection of data records or specimens – for example, a species observation record or soil sample – and *data analysis* to refer to the generation of digital data (known as metadata) to categorise, define or give additional details on a digital asset such as an image, video or text file. Strictly speaking, both of these processes would be described as data gathering by both Haklay and Wiggins and Crowston, but data collection projects would be *investigative*, while data ana-

[1] See chapter 4 for an in-depth discussion of VCS literature.

[2] Problem solving is often part of data collection and data analysis initiatives; see, for example, Curtis (2014) and Lintott et al. (2008)

lysis projects would be *virtual* according to the typology defined by Wiggins and Crowston (2011).

Data collection tasks require citizens to gather data or materials such as images to enable scientific research, and cover tasks such as observation, measurement and sample collection (Wiggins and Crowston, 2012). Unlike in more traditional CS approaches, volunteers receive training through technology-mediated methods, such as tutorials, videos or technology-mediated communication with professional scientists or peers, either in addition to or in place of face-to-face training (Catlin-Groves, 2012; Starr et al., 2014). Data collection may include the submission of photos, recordings, textual reports, GPS data, the use of specially designed apps or the completion of surveys, among other tasks. Due to the nature of data collection projects, volunteers generally carry out offline tasks in real-world locations, before using web-based technology to submit data (Wiggins and Crowston, 2011).

The VCS project *eBird* is an example of a data collection project. Launched in 2002, eBird is one of the largest VCS projects (Yu et al., 2010), with over 500,000 volunteers (Price and Dorcas, 2011). eBird leverages the hobby of amateur birdwatchers, encouraging volunteers to submit their observations to an online database and, in doing so, aid scientific research (Wood et al., 2011; Yu et al., 2010). Volunteers within eBird contribute vast amounts of data, with 1.5 million contributions from within North America in 2010 (Yu et al., 2010), with 92,000 observations submitted each day (O'Donnell and Durso, 2014). After an initial focus on North America, the project has grown to include observations from all over the world (Catlin-Groves, 2012). Due to the visualisations available within the project (see Figure 1), volunteers benefit from the increased accessibility of bird watching data and thus, as well as having a strong offline component, the online component further enables participant activity (Catlin-Groves, 2012). Observations submitted to eBird are used in a range of fields, including statistics, biology and computer science (Roy et al., 2012).

In contrast, *data analysis* (or *virtual*) tasks make use of human computation to produce metadata as a means of classifying or describing data or materials such as images for further analysis and research. Analysis tasks typically include vast numbers of assets and require multiple analysis submissions for each subject as a measure of quality control (Wiggins and He, 2016). As a result, such initiatives must appeal to large numbers of volunteers (Tinati et al., 2015b). Unlike data collection projects, data analysis projects tend to be loc-

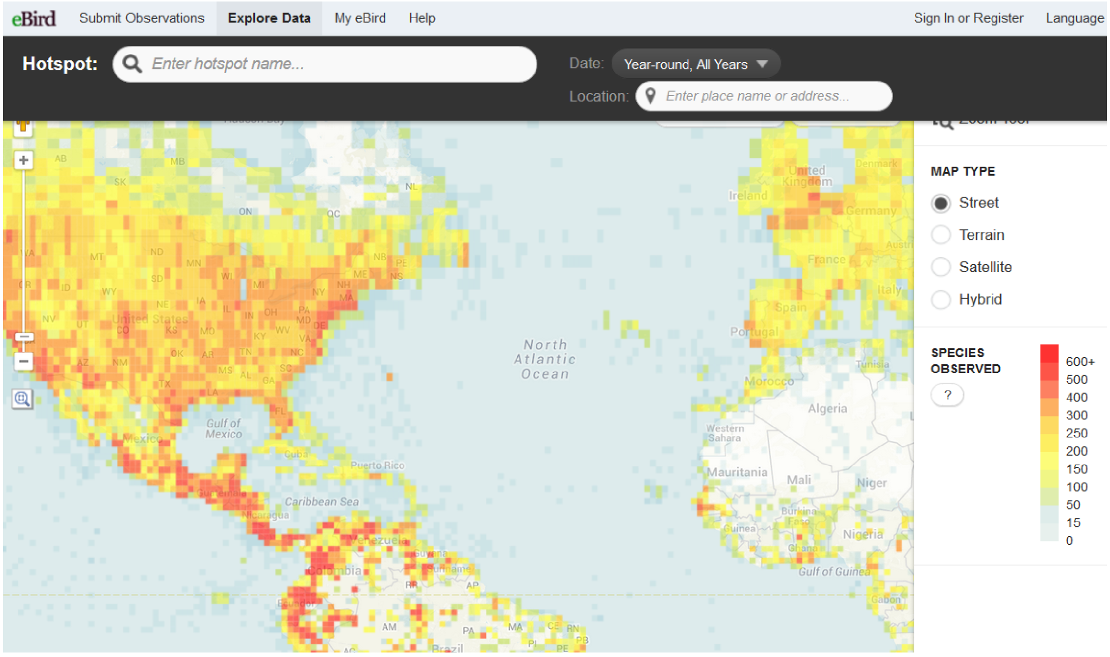


Figure 1: Map visualising observations recorded within the eBird project

ation independent, being instead rooted in a virtual space from which submissions are made. Tasks vary between fields of study, projects and platforms and there is no dominant typology or categorisation for data analysis processes, but examples include:

- **Cataloguing**: Classifying items by characteristic or content, as in the Snapshot Serengeti project where volunteers classify images according to the animal species they represent (Swanson et al., 2015).
- **Mapping**: Mapping structures within images, as in the CosmoQuest platform, where volunteers map and trace structures on images from the surface of planets and asteroids (Gugliucci et al., 2014).
- **Transcribing**: Digitising textual data from images or physical records, as in the Old Weather project where volunteers digitise weather and location readings from ship logbooks (Roy et al., 2012).

Galaxy Zoo is arguably one of the most well-known data analysis VCS projects (Wiggins and Crowston, 2011). Drawing on images from the Sloan Digital Sky Survey, the project aims to generate morphological classifications of galaxies and replaces more traditional workflows where such classifications were completed by small teams or individual PhD students (Galloway et al., 2015;

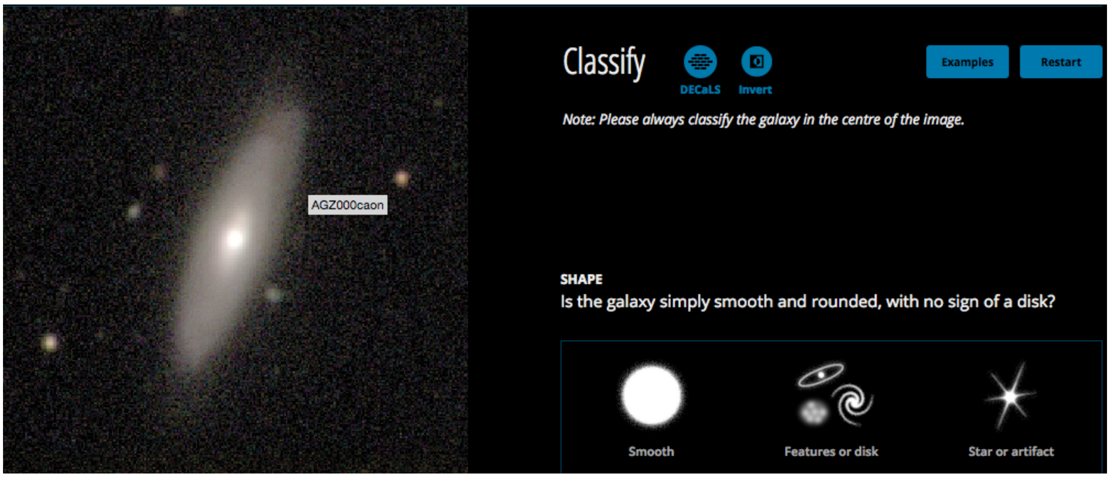


Figure 2: Galaxy Zoo classification interface

Lintott et al., 2008). First launched in 2007, initial early modest estimations for classifications were quickly surpassed, leading to the development of Galaxy Zoo 2 (Willett et al., 2013). Now in its fourth incarnation, Galaxy Zoo has generated millions of classifications across 1.5 million assets from over 300,000 volunteers (Tinati et al., 2015b). Users' classifications represent only part of the research process, which makes use of more traditional, scientist-led methods to analyse data from the classified images (Madison, 2014). Although two high-profile serendipitous community-led scientific discoveries within Galaxy Zoo have led to publications with citizen scientist co-authors[3], the project task is relatively restrictive, and Galaxy Zoo thus corresponds to the *distributed intelligence/contributory* category described by Haklay (2013) and Wiggins and Crowston (2012).

This distinction between data collection and data analysis – or investigative and virtual projects – while potentially clear-cut in many cases, is arguably unsuitable for describing the current Citizen Science landscape. This is in part due to the existence of projects such as Tiger Nation[4] and iSpot Nature[5] which combine data gathering tasks with metadata generating 'data analysis' tasks (Mason et al., 2012; Silvertown et al., 2015). In addition, and as demonstrated by eBird, even explicitly offline projects and activities have moved increasingly online as the use of the web has grown, offering training materials, guidance or

[3]Galaxy Zoo has led to two high profile published discoveries: a class of compact extremely star-forming galaxies (Cardamone et al., 2009), and 'Hanny's Voorwerp', a quasar light echo (Lintott et al., 2009).

[4]Tiger Nation CS project – `https://www.tigernation.org/` [Accessed 30 Aug 2018].

[5]iSpot Nature project – `https://www.ispotnature.org/` [Accessed 30 Aug 2018].

Source	Basis	Categories
Tinati et al. (2015b)	Task type	Data collection, Data analysis, Problem solving
Wiggins and Crowston (2012)	Participant involvement	Contributory, collaborative, co-created
Haklay (2013)	Participant involvement	Crowdsourcing, distributed intelligence, participatory science, extreme Citizen Science
Wiggins and Crowston (2011)	Project purpose	Action, Conservation, Investigation, Education, Virtual
Various (e.g. Kullenberg and Kasperowski (2016))	Project discipline	Varied, particularly in terms of specificity – examples include: Astrophysics, Ecology, Humanities, Immunology

Table 1: Citizen Science classification typologies

opportunities for communication and feedback through virtual portals and tools (Catlin-Groves, 2012). Similarly, with the increasing use of smartphone devices and the rise of smartphone-based apps, the distinction between virtual and offline data collection activities has become increasingly blurred (Newman et al., 2012). For example, the mySoil app enables participants to share their knowledge through Citizen Science, without necessarily requiring that they travel to a location to gather data (Roy et al., 2012). The boundaries of Citizen Science varieties are likely to become further muddied as new technologies permeate scientific research, offering new ways for participants to interact with scientists, one another and the research landscape. At the same time, this dichotomy persists in the way Citizen Science is viewed as an approach, as Wynn (2017) argues that it is only technology-mediated Citizen Science – that is *virtual* Citizen Science – that has been accepted as a valid research methodology by the scientific community.

2.2 Online Communities

Online communities are web-based environments in which people gather to work towards common goals, socialise, share knowledge and resources, or simply communicate (Kraut et al., 2012). Such communities vary in size and may take diverse forms – from simple message board forums to complex, immersive and social virtual worlds such as *Second Life* and *World of Warcraft* – although the majority of online communities take the form of textual discussion forums or email groups (Lampe and Johnston, 2005; Kraut et al., 2012). Similar to Citizen Science, many of these formats started as offline technologies, before being reinvented on the web (Preece, 2001).

The most common reasons for participation in online communities include exchanging and accessing information, developing social support networks and creating friendships (Ridings and Gefen, 2004). As a result, it is not just the content that community participants value, but the opportunity to interact with fellow contributors and access the underlying social infrastructure of the platform as well. These findings persist over a range of domains and platforms, from social networks to commercial services (Ridings and Gefen, 2004).

In spite of the diversity of online communities, the underlying purposes of many systems overlap or are at least very similar. In fact, Armstrong and Hagel (2000) propose a typology of just 4 types of online community: communities of *transaction*, *interest*, *fantasy* and *relationship*. In communities of transaction, participants buy and sell goods and services, with opportunities to provide and seek information before making transactions. Communities of interest are arguably the most typical form that online communities take, with participants interacting, discussing and sharing opinions and materials with one another – generally around specific subject areas. Communities of fantasy are similar to the other types of community, but are distinguished by the behaviour of participants, who take on specific imaginary roles and profiles, imitating existing real-world and fictional figures[6]. Finally, in a community of relationship, participants support one another with more personal issues such as sufferers of diseases, or victims of crimes. These categories need not be mutually exclusive and in fact, Armstrong and Hagel suggest that the most effective communities, at least in commercial spaces, should aim to combine these categories.

Based on this model, online Citizen Science projects can be seen as examples of online communities of interest, but also of transaction. In this particular case, the transaction taking place is not of money or goods and only partly an exchange of information between participants. Instead, in crowdsourcing, project managers approach an online community (the 'crowd') who each contribute a portion of their knowledge, building up a large final understanding of a problem (Brabham, 2013). Still, the most fundamental aspects of the com-

[6]Although Armstrong and Hagel (2000) describe communities of fantasy in terms of conscious role-play, these may also describe activities such as MMORPGs where the participant adopts a role as an in-game character, rather than through his or her behaviour (Kraut et al., 2012).

munity of interest and of online communities are represented in crowdsourcing platforms. Interested individuals (predominantly experts and professionals) self-select to contribute and share their knowledge with others (Brabham, 2013).

How best to design online communities is a question that has been studied using various frameworks and analytical methods. Ren et al. (2007) analysed theories of commitment with regard to online communities in order to identify design decisions which may lead to greater commitment. The authors demonstrated that specific design decisions such as constraining or encouraging discussion among the community can influence group formation and resulting commitments between community members and in doing so, influence the form that community participation takes.

Preece (2001) explored the dimensions of sociability and usability in an effort to explore the concept of success in online communities. The resulting framework considers facets of sociability: volume of participation, reciprocity in contributions and benefits, quality of contributions, and participant behaviour. Usability dimensions are also key components of the framework: ease of use, speed of learning, measures of productivity, and user retention. These dimensions are of particular interest given the objective of this thesis to identify features associated with online communities and contributions in VCS projects.

Similarly, Iriberri and Leroy (2009) analysed online communities within the framework of information systems life cycles to further explore the concept of success in online communities. Dimensions evolve throughout the cyclical framework, from conception and purpose, to ensuring security and reliability during the creation process. Quality assurance and encouraging interaction become important during the growth phase, while mature communities must further focus on rewarding and encouraging interactions through events. Such a framework demonstrates the evolving nature of success and highlights the importance of early design decisions on later project outcomes.

Kraut et al. (2012) studied a diverse body of communities, ranging from simple, crowdsourcing efforts such as *CAPTCHA* and *Mechanical Turk* to more complex communities including MMOs such as *World of Warcraft*. Each of these communities was analysed based on empirical observations informed by key theories from the social sciences. The authors devised a total of 175 design claims, across five areas: encouraging contribution, encouraging commitment, regulating behaviour, dealing with newcomers and starting new communities. These design claims provide evidence-based guidelines for assessing online community success.

With regard to gamification, Mekler et al. (2013a) explored the impact of points and contextual framing of tasks on volunteers' intrinsic motivations and performance in an image annotation task. Points were found to increase the quantity of tags generated, while framing had no significant effect on quantity. A combination of points and contextual framing was shown to have a significant effect on the time spent per tag, compared to just points or framing alone. Framing was associated with an increase in tag quality, an effect which was not seen with points. These findings suggest interface features can impact participant engagement and the effort expended by participants.

2.3 Collective Action

Collective action is the theory that groups of individuals with common interests or needs will act not as individuals, but rather as a *collective* to fulfil those needs (Olson, 1965). In other words, in a collective action context, each participant puts the interests of the entire group or community before his or her own needs and acts not only for his or her own benefit, but first and foremost for the benefit of the entire group or community. In so doing, the participant benefits as part of the collective. The theory has been applied to explain engagement in human computation initiatives, but may also describe simple community tasks, such as building a barn (Marwell and Oliver, 1993; Michelucci, 2013). Collective action generally presupposes some kind of social ties exist between members of the group, but at its heart, it is defined simply in terms of the existence of common interests among group members and the ultimate outcome of communal benefits (Marwell et al., 1988; Marwell and Oliver, 1993).

Issues have been raised with collective action as a theory, given that participation patterns are asymmetric, with few participants contributing the majority of work, so that all may benefit – in essence, an issue of individuals versus groups (Hardin, 1971). This is best illustrated by the *social* or *prisoner's* dilemma problem, in which each participant faces the choice between contributing and receiving the common benefit, or not having to contribute but still receiving the common benefit when the action is completed (Ostrom, 1998). Game theory suggests that all participants will make the choice to not contribute, so that they might achieve the maximum reward with the minimum effort (Ostrom, 1998). As a result, the work will of course not be done and ultimately, nobody will benefit. To express this theory in simpler terms, Olson (1965) posited the *zero contribution thesis*, which states that "rational, self-interested individuals will not act to achieve their common or group interests," except in the event of external factors such as coercion, or where the number of individuals involved is small.

To counter these criticisms, collective action has been viewed in the light of *social identity theory*[7], with the implication that individuals engage in collective action to develop their own social and activist identities (Van Zomeren et al., 2004). Furthermore, these issues have been suggested by some observers to *justify* collective action in some contexts, with activists and individuals engaging in – or at least commencing – collective action under the assumption that others will not do so and thus, tasks will otherwise not be completed (Oliver, 1984). In any instance of collective action, the resulting output from collective action is not a result of a large community of participants, but rather a small number of highly dedicated, active individuals (Marwell and Oliver, 1993). Even so, collective action requires a certain level of critical mass, in that without a sufficiently large community with sufficiently invested and active individuals, any resulting output is unlikely to be successful – a factor that has also been observed within VCS (Marwell and Oliver, 1993; Tinati et al., 2015b). As a result, for any collective action initiative to be successful, it must rely on initial actors to be altruistic or to at least recognise, as Oliver (1984) suggests, that if they are not willing to act, then nobody will.

2.4 Gamification

Gamification describes the use of "game elements in non-gaming contexts" (Deterding et al., 2011). The aim of gamification is to motivate participation in an activity, by leveraging the naturally motivating characteristics of video games and associated elements and extend these to non-gaming tasks (Sailer et al., 2017). Gamification has been employed in a wide array of contexts, from education and research to business and marketing (Seaborn and Fels, 2015) and a significant focus of gamification research concerns evaluating the impact of diverse game elements in varying contexts and platforms (Jia et al., 2017; Kumar, 2013; Mekler et al., 2013b). Specific gamification features depend on the con-

[7]The theory that "people generally strive for and benefit from positive social identities associated with their membership groups" (Van Zomeren et al., 2004).

text and platform type to which they are applied, but Seaborn and Fels (2015) describe the following common gamification elements:

Points A reward mechanism, which allows players to accumulate a numerical score as they contribute, to compete with their own previous performance or that of other players. Points also allow for users to track their progress.

Badges A reward mechanism. Icons displaying previous achievements and progression.

Leaderboards A visual display of all or a subset of players' progression (for example, point score) or achievements.

Progression Inclusion of milestones which players must reach or surpass – for example, completing levels or levelling up.

Status Features such as ranks or titles, which convey status by being displayed and made visible to other players.

Levels Different environments, tasks, etc., which become increasingly more difficult as players proceed from one level to another.

Rewards An object such as a tangible prize, which players desire and can gain through participation in the game.

Roles A character element which a player may take on as they contribute to the game or task.

In discussing the application of gamification to VCS, it is essential to make the distinction between the use of game elements and *Games with a Purpose*, where participants carry out human computation tasks as part of a casual game experience. (Siu et al., 2014). While human computation initiatives such as Citizen Science projects may make use of gamified elements (such as points and rewards) to encourage player engagement with tasks, it is the diegetic, narrative-driven and immersive nature of the elements in Games with a Purpose which sets them apart from other projects (Prestopnik et al., 2014).

Launched in 2008, FoldIt is an example of a GWAP, where volunteers complete puzzles to identify the process for folding proteins – a task which is difficult to accomplish with more traditional computation methods, as current algorithms are inaccurate and computationally expensive (Curtis, 2014; Khatib

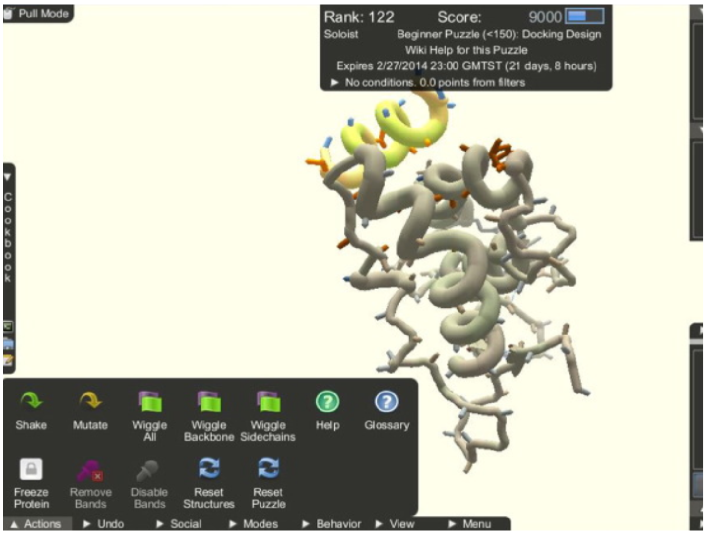


Figure 3: FoldIt puzzle interface adapted from Curtis (2014)

et al., 2011). Players' solutions to protein folding puzzles are rated, based on the energy required to produce proteins and the similarity of these solutions to natural protein configurations, with olutions awarded points according to these ratings (Curtis, 2015a, 2018b). The FoldIt game is online and multiplayer, as participants compete to achieve the highest score and communicate through integrated and external discussion platforms (Khatib et al., 2011). Volunteers may earn points (and thus compete) individually or as a team (Curtis, 2014). Players are encouraged to collaborate and discuss solutions with other volunteers through a live chat function, with the opportunity to use asynchronous communication platforms (Curtis, 2015a). Due to the complex nature of the protein folding process, FoldIt is a relatively complex game when compared with other online multiplayer games and players are required to complete a lengthy process of 32 tutorial puzzles before reaching assessed puzzles (Curtis, 2015a). Despite this difficulty, players have contributed to the development of protein folding algorithms surpassing previously published algorithms (Khatib et al., 2011).

A diverse set of studies have considered the role that gamification and games play in motivating participation in a number of Games with a Purpose and VCS projects (Bowser et al., 2013; Eveleigh et al., 2013; Iacovides et al., 2013). Results from such findings have been mixed, with players identifying both positive and negative aspects of gamification in VCS. Quantitative analyses have suggested that self-reported motivations do not align with volunteer contributions (Mekler et al., 2013a). These findings are further discussed in chapter 4.

Additional studies have considered the application of gamification to crowdsourcing activities. Jia et al. (2016) explored the impact of individual gamification elements on participants according to personality characteristics. In crowdsourcing mobile data collection tasks, Dergousoff and Mandryk (2015) found that the use of freemium model based in-game rewards increases engagement, but such models lower the quality of submissions. In team-based task completion, partnering with highly effective team-mates has been found to reduce players' own intrinsic motivations and perceptions of tasks (Luu and Narayan, 2017). Brouwer (2016) conducted a survey of professional work teams in the Netherlands around the subject of intra-team competition and found that such competition has both negative and positive effects on team performance, through increased task complexity but reduced psychological safety. Zheng et al. (2011) analysed motivations for participating in crowdsourcing competitions in paid microtask work, showing that engagement is linked to the nature of contest demands. In particular, the authors found that intrinsic motivation and associated contest-design decisions were key to driving participation in competitions, but nevertheless must be balanced with extrinsic factors. Similarly, introducing cognitively demanding challenges to games has been found to increase player activity, while physically challenging tasks do not (Cox et al., 2012).

More specific to Games with a Purpose, but in a non-Citizen Science context, Siu et al. (2014) explored the impact of competition-based scoring and reward mechanisms compared with collaboration-based reward mechanisms. Harteveld et al. (2016) created and tested a digital game to allow non-scientists to engage in scientific experimentation, demonstrating that players require the support of AI and task structuring to get the most out of the game.

2.5 Crowdsourcing

One range of research perspectives on Citizen Science considers Virtual Citizen Science projects as crowdsourced, human computation projects. Crowdsourcing and human computation are related concepts, in that both concepts refer to outsourcing a task to an individual or group of individuals. However, human computation exclusively refers to outsourcing a task from a computer to a human individual, while crowdsourcing refers to outsourcing a task from an individual worker to a larger group of individuals (or 'crowd') (Quinn and Bederson, 2011).

Despite this distinction, an overlap exists: any context in which the crowd of individuals may serve as a replacement for either a computer or human agent can be considered an example of both human computation *and* crowdsourcing (Quinn and Bederson, 2011). Citizen Science exists within this space – while the tasks involved are difficult for humans, they can also not be carried out using methods involving traditional workers, due to the large volume of data which such workers would have to deal with (Raddick et al., 2009a).

Human Computation, as noted, refers to the use of humans as replacements for computational processes and technologies and is generally employed in the case of tasks which computers cannot perform, or which are difficult to solve through ordinary computation (Quinn and Bederson, 2011). Little et al. (2010) describe two task categories which necessitate human computation over more traditional methods: creative tasks, where the user must produce new content, such as a description, image or concept, and decision tasks where a user must take multiple examples of created content and rate those according to their appropriateness for a given context. Quinn and Bederson (2011) note further ways in which tasks may include other human elements, such as the need for specific (human) skills such as command of a given language in order to understand content or certain qualifying characteristics such as being an expert in a given area, without which tasks would not be possible. A more unique form of human computation can be found in the aforementioned "Games with a Purpose", which make use of a human capacity and desire for play to achieve tasks (Jain and Parkes, 2009).

Crowdsourcing, while similar, does not necessarily involve any attempt to replace a computer participant with a human one. Howe notes that crowdsourcing systems may serve to improve the experience of their participants through users' attempts to teach each other and to improve other users' task completion capacity (Howe, 2008). On the other hand, it is clear that an overlap does exist – in keeping with this overlap between human computation and crowdsourcing, decision making, creation and the use of specific skills all also correspond to crowdsourcing tasks described by Dunn and Hedges (2013), alongside many skills which may replace a computer with a human participant.

Perhaps unsurprisingly, given the large number of participants involved, crowdsourcing projects are seen to be inherently collaborative. Howe (2008) describes crowdsourcing as enabling "unprecedented levels of collaboration", both between the users of the system and between the users and hosts of the system. Pan and Blevis (2011) surveyed literature surrounding crowdsourcing to identify the implications of crowdsourcing for collaboration in three areas: academic, enterprise and social, concluding that the collaborative aspect of crowdsourcing added particular forms of value in each of the three categories, such as allowing for innovation in academia, or leading to greater profits in enterprise. Afuah and Tucci (2012) describe a specific subset of crowdsourcing initiatives as "collaboration-based crowdsourcing", where smaller groups within the community may form teams and attempt to solve problems. This aligns with the types of crowd-sourcing task identified by Dunn and Hedges (2013), some of which explicitly involve collaboration such as collaborative tagging tasks or tasks involving translation, where users must generally collaborate to produce a coherent translation.

While human computation is generally more variable, it too is by nature collaborative. As Quinn and Bederson (2009) state, any human computation initiative uses networks of humans to complete tasks which could not be completed by computers or small groups of humans. In addition, given the overlap between human computation and crowdsourcing, any human computation task that can also be considered an example of a crowdsourcing task can be considered to be collaborative. There is also a growing literature surrounding human computation tasks which discusses the emerging concept of human-machine collaboration, such as in the case of Hu et al. (2010). However, this is largely beyond the scope of this research, for the purposes of which the underlying technology will be seen as *supporting* collaboration between the human elements of the system, rather than being an active collaborator (see Social Machines).

2.6 Social Machines

In his book *Weaving the Web*, Berners-Lee (1999) describes "processes in which the people do the creative work and the machine does the administration,"[8] which he calls Social Machines. Such processes encompass a number of goals such as human computation, social computing and collective intelligence, to name but a few (Shadbolt et al., 2013). Furthermore, these processes align to a number of web-based technologies used by millions of users, such as social

[8] In the next line, Berners-Lee (1999) goes on to describe electronic voting processes as an example of a social machine and so "creative" here refers to a potentially simple process such as decision making, rather than the creation of a work of art, music or literature.

networking sites like Facebook and Flickr, crowdsourcing platforms such as Ushahidi and collaborative platforms like Wikipedia (Smart and Shadbolt, 2014). Citizen Science projects share such processes and are explicitly described in the literature to be examples of social machines[9].

While key processes from Citizen Science such as human computation are potential components of social machines, it should be noted that opportunities for discussion interaction are equally important aspects of social machines in themselves. Buregio et al. (2013) note that many Web 2.0 services such as blogs and social networking sites can in fact be considered early examples of social machines and indeed, Berners-Lee (1999) described early discussion systems in similar terms. However, the levels of interaction enabled may vary strongly between systems and the term *social* in social machine does not necessarily entail any kind of social or discussion-based process.

In fact, Smart et al. (2014) describe eleven variables which together serve as a taxonomy of social machines. Particularly valuable to this research are the consideration of *sociality* – the extent to which interaction is supported by the system – *inter-dependence of user contributions* and *visibility of user contributions*, which describes whether users are able to view each other's contributions. Other factors described include the *size* of the community, the *extent* to which the whole community of users contributes and the *diversity* of tasks which users can engage in.

Similar classifiers are identified by Vass and Munson (2015), who consider the factors which make up a 'profile' of social machines. This is of particular interest for this research in terms of the agency, symmetry and reflexivity assigned to human agents – that is, the extent to which individuals within the system are able to feedback and reorient the goals of the wider projects. In terms of symmetry – the level at which all human agents within the machine have the capacity to influence the outcomes and processes of the machine itself – the more symmetrical, the more equality there is between human agents. Related to this are the concepts of recognition and responsivity, essential aspects of reflexivity. Here, recognition describes the agency assigned to individual agents within the wider system – are participants able to construct their own identities, or do a subset of agents hold the power and responsibilities? Similarly, responsivity describes the effectiveness of the system in responding to the activity of the constituent agents within the system.

[9]see for example: De Roure et al. (2015); Smart and Shadbolt (2014)

These measures hold interesting parallels with the typologies outlined by Haklay and by Wiggins and Crowston. Rather than Citizen Science projects corresponding to one form of social machine, instead they can be considered as occupying a range of positions in terms of symmetry, recognition, responsivity and agency more broadly. Indeed, at their lowest level as described by Haklay, projects can be seen as highly asymmetrical with low recognition and low responsivity – scientists set tasks and dictate process, rules and carry out many of the internal processes. Volunteers, conversely, gather data from sensors and have little opportunity to dictate the terms of the project (Haklay, 2013). At the other end of the spectrum would be the co-created projects where volunteers are involved at all stages of the scientific process, defining questions, designing studies and participating in the gathering and analysis of samples (Wiggins and Crowston, 2012). It becomes clear, then, that projects with the lowest reflexivity are unlikely to involve community interaction at any significant level and such interaction is highly unlikely to influence the outcomes and success of a project.

2.7 Distributed Cognition

Distributed Cognition as a theory does not necessarily cover collaborative actions and may simply cover an individual's use of a tool or tools (Rogers, 2012). In this way, distributed cognition overlaps somewhat with ANT, with both human and machine agents, each of equal importance to the overall system and, in the case of DC, flow of information. In fact, distribution cognition analysis will typically consider all types of communication taking place – both verbal and non-verbal – as well as the "coordinating mechanisms" such as rules and the way in which information is accessed – i.e., the use of tools and technologies within the system (Rogers, 2012).

Distributed Cognition theory has been applied to VCS and related initiatives in specific contexts. Watson and Floridi (2018) describe it as a core function of the serendipitous discoveries that have taken place within Zooniverse through the socio-technical features – predominantly the discussion platform *Talk* – as participants share knowledge with one another and the science team and in doing so, gradually build up an understanding of the underlying discoveries. Smart (2017) extends this view to the inherent characteristics of crowdsourced initiatives such as human computation, of which VCS is but an example, noting that there is a common assumption when discussing these initiatives, that information is produced from the contributions of multiple human agents through an extended or distributed cognitive process. This view was similarly described by Sakamoto et al. (2011), in noting that crowdsourcing as a methodology may be employed by researchers to examine the processes involved in Distributed Cognition.

Yet at the same time, there are questions regarding the extent to which VCS fits within this paradigm. Taking again the perspective of VCS as a Social Machine, information flows in VCS may be highly restricted by the machine agents, with each participant seeing little or nothing of other participants' contributions[10] (Mugar et al., 2014). If we were to consider only *Socially Distributed Cognition* – the transmission of information between members of a group or organisation (Hollan et al., 2000) – then information within VCS tasks could not be said to be distributed in this sense, as no information is transmitted between participants based on task alone. Even when accounting for machine agents Hollan et al. (2000) talk of a conscious process by which individuals pass cognitive tasks to machine agents to reduce their own workload, but this does not align to the crowdsourced tasks described by Smart (2017). Rogers and Ellis (1994) describe such distributed cognition as a somewhat delayed process, as a person uses one tool – for example, word processing software – but must then wait for information to be fed back and in turn, offloads more tasks to additional tools. But this is arguably not the case in a crowdsourcing service, or at least, certainly not in the design requirements described by Mugar et al. (2014), where there is no information to be fed back at all. Some have even questioned the extent to which non-human agents can engage in cognition, which they view as a very human process (Rogers, 2012). Nardi (2002) describes such a view as an error and notes a lack of specific features which should be sought in carrying out distributed cognition analysis.

2.8 Actor-Network Theory

Actor Network Theory describes social processes as networks of various actors – the *actor network* – with common interests or goals, in which the outcome is a factor of the various actors within the system (Walsham, 1997) It represents the view that social structures and relationships are inherently influenced

[10]This theme is revisited and discussed in more detail in chapters 4 and 5

by the contexts in which they occur and that as researchers, we cannot and should not make empirical observations without consideration for these contexts (Whittle and Spicer, 2008). In ANT, almost any constituent part of a system or network can be viewed as an actor – from machines to weather phenomenon, to bacteria – and there is no distinction made between the capacity of human and non-human actors to have agency within and influence upon the process or system (Latour, 2004). This does not imply some form of artificial intelligence or that there is equality between human and non-human actors – clearly any human actor has a capacity and the freedom to act in a way that non-human actors do not (Sayes, 2014). Instead, ANT represents the recognition of other forms of agency, a recognition, as Latour (2004) describes, that these non-human actors may be able to "authorize, allow, afford, encourage, permit, suggest, influence, block, render possible, forbid," or to otherwise facilitate or impede the actions of human actors. In more simple terms, (Bencherki, 2017) describes the role ascribed to non-human actors as a capacity to "make a difference" in the specific process.

ANT offers valuable perspectives for the study of technology systems, in which it is often difficult to extract the social and human processes from the technology in which those processes occur (Walsham, 1997; Tatnall, 2005). The theory should not be viewed as exploring and building a network of every single actor within the network, but rather ANT represents the acknowledgement that there is a level of uncertainty involved in understanding the processes and factors involved in any given system. This is particularly valuable for studying the system, lest as Law (1992) states, "we might start with interaction and assume that interaction is all that there is". Instead of attempting to distinguish between interlinked human and non-human actors – for example, technology – an ANT perspective recognises and is receptive to the possibility that an outcome may be influenced by multiple agents, which cannot be removed from the context in which they occur (Bencherki, 2017; Tatnall, 2005).

Chapter 3

Methodological Approach

This chapter details the philosophical and methodological approach which underpins the studies described in chapters 4 to 8. Firstly, the philosophical paradigm of *pragmatism* is introduced, along with a description of *mixed methods* popularly associated with pragmatist philosophy and research. This is followed by an explanation of the design of the experimental and empirical approach within this thesis, as well as an explanation and justification of the projects sampled as a basis for the research methods in the following chapters. In section 3.5, the qualitative methods employed within chapters 4, 5 and 6 are presented – specifically the literature review, project review and interview approaches. This is further followed by an explanation of potential quantitative methods, as well as the justification and reasoning behind the specific methods used in chapters 7 and 8.

3.1 Mixed Methods

Mixed methods research "combines elements of qualitative and quantitative research approaches," making extensive use of the integration of both methods and results as a "core component," of the process (Bazeley, 2017). Both qualitative and quantitative methods are varied and complex and even the absolute most extreme and 'pure' cases of either approach may include aspects or characteristics of the other (Tashakkori and Teddlie, 2010). However, mixed research and methodologies go beyond this, representing an active rejection of the incompatibility of methods thesis[1] and in doing so, mixed methods studies represent a philosophically pragmatist approach that does not correspond to the postpositivist[2] and constructivist[3] paradigms often associated with qualitative and quantitative research methods (Tashakkori and Teddlie, 2010). Rather than adopting or rigidly adhering to a given paradigm, mixed methods approaches are more adaptive and open to "what works", with researchers tailoring their procedures based on what is deemed best for the research topic and questions at hand (Yanchar and Williams, 2006).

On the other hand, this does not suggest that quantitative and qualitative concepts are mixed together haphazardly or without thought. On the contrary, mixed methods research tends to follow one of six broad methodological approaches (Cresswell and Plano Clark, 2018; Cresswell, 2014). In parallel or convergent mixed methods studies, two or more quantitative and qualitative studies are carried out distinctly before integrating and triangulating the results to build a broader understanding of the overarching research question (Bazeley, 2017; Cresswell, 2014). Alternatively, such research may be *sequential*, with the results of one study explained by and informing the development of a second study. In explanatory sequential mixed methods research, the initial research is quantitative and these findings are explained and expanded on through qualitative approaches – for example, survey results may be explained through more focused and more detailed interview sessions (Cresswell, 2014; Tashakkori and Teddlie, 2010). Alternatively, sequential MM research may be exploratory, with an initial qualitative process that leads to the development of a model, framework or typology that is then used in the design of a subsequent quantitative study (Cresswell, 2014; Tashakkori and Teddlie, 2010).

In more complex cases, a qualitative or quantitative study may be embedded in an existing methodological process or study, to provide a greater breadth of interpretation (Cresswell, 2014; Cresswell and Plano Clark, 2018). In transformative mixed methods, any one of these approaches may be employed for

[1]The incompatibility of methods thesis posits that "compatibility between quantitative and qualitative methods is merely apparent" with "hidden epistemological difficulties" between the two approaches (Howe, 1988).

[2]Postpositivist research is associated with identifying effects and their associated causes and is generally associated with strict experimental methodologies and thus, quantitative methods (Cresswell, 2014)

[3]Constructivist research takes the view that the researcher develops meaning and understanding through the research process and thus, that no observation or experimental outcome can be completely accurate or pure. It is therefore associated with qualitative methods and is, in some ways, the opposite of the postpositivist approach (Teddlie and Tashakkori, 2009).

the purposes of social justice or in the context of social theories such as feminist or queer theory, among others (Teddlie and Tashakkori, 2009). In contrast to distinct parallel and convergent methods, quantitative and qualitative processes may inform one another to enable a case study of a given context, community or phenomenon (Cresswell and Plano Clark, 2018). Finally, in multiphase mixed methods, multiple quantitative and qualitative studies may be conducted in sequence with the design of each study being highly dependent on the results of previous studies within the sequence (Cresswell, 2014).

This thesis makes use of a multiphase mixed methods approach, where the outcomes of each chapter inform the methodology and approach of the subsequent chapters (Cresswell, 2014). In writing about this study design, Cresswell and Plano Clark (2018) noted that as a definition it had become too general to be of much use and so it is necessary to discuss in more detail how this process occurs, as at heart any mixed method study is multiphase. Each of the studies conducted in this thesis is sequential and the outcomes, issues and questions raised by each study are an essential element in designing and informing the subsequent studies. In turn, study two and three form an explanatory sequential design, where the results from the quantitative observation findings from study two are explained and interpreted in light of the qualitative interview findings from study three (Tashakkori and Teddlie, 2010). Similarly, studies three and four in turn form a mixed methods case study (Cresswell, 2014), where the qualitative interview findings and quantitative data analysis findings serve as a case study of task sociality through competitions in EyeWire. An overview of the process can be seen in figure 4.

3.2 Pragmatism

Pragmatism is the most popular philosophical and epistemological paradigm within mixed methods research (Teddlie and Tashakkori, 2009). In terms of philosophy, Peirce (1905) was the first to define the paradigm of pragmatism, describing it as a method "to trace out in the imagination the conceivable practical consequences [...] of the affirmation or denial" of a given concept (Cherryholmes, 1992). Pragmatists hold the view that while true statements share a common characteristic that may be known as 'truth', the nature of this objective truth and the concept of truth itself is not of great interest or value (Rorty, 1982). While there may be an objective world external to that perceived

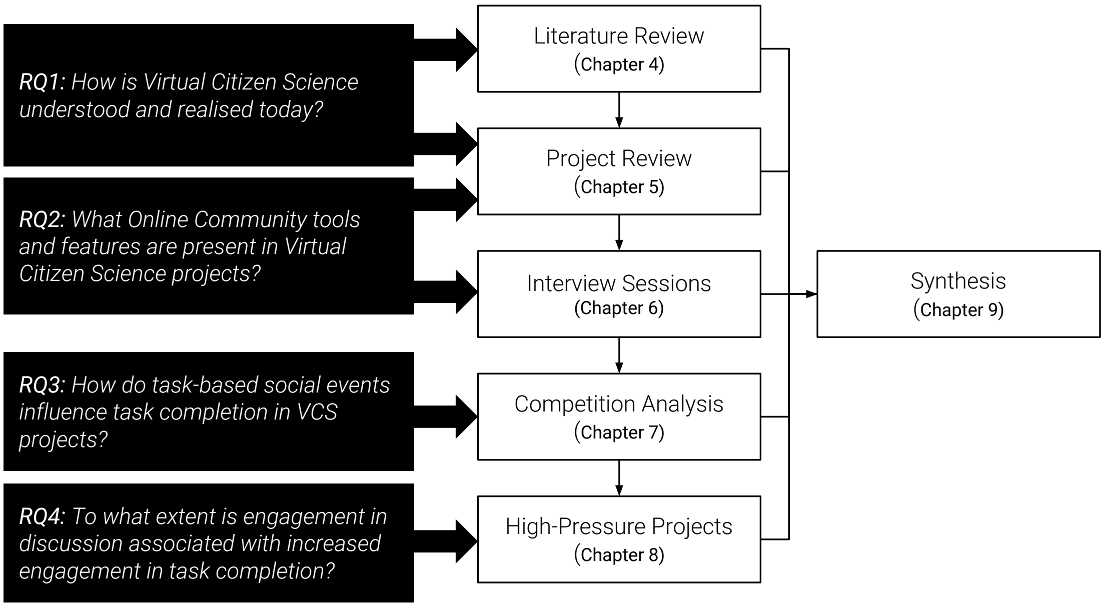


Figure 4: Thesis overview, showing how each study builds on previous studies to create an overall synthesis of findings.

by individuals, it is of no great importance to question the nature of reality and there are more interesting and pressing subjects to consider (Rorty, 1982; Cherryholmes, 1992). In essence, pragmatism is concerned with actions and consequences, rather than the theoretical underpinnings and prerequisites of those actions (Cherryholmes, 1992). It is of greater consequence to a pragmatist that an action has a given outcome than questioning why this may be.

When it comes to research and methodologies, pragmatist perspectives are predominantly focused on the application of problems and issues and developing responses to these problems based on what works best for a given issue or context (Cresswell, 2014; Yanchar and Williams, 2006). Johnson and Onwuegbuzie (2004) simplify the pragmatist perspective as it relates to the research process in stating that "if two ontological positions...do not make a difference in how we conduct our research" then for all intents and purposes, the difference between the two positions does not matter for the sake of the research process. Morgan (2007) describes pragmatism in terms of a rejection of the previous view of science and research as constrained to paradigms, a view most prominently held by Kuhn (1970), but also adopted by others. At its most simplest, a pragmatist approach to research can be seen as adopting and embracing solutions and methods that work and abandoning those that do not (Howe, 1988; Tashakkori and Teddlie, 2010; Yanchar and Williams, 2006).

There are also those authors who view pragmatism not necessarily as a philosophical paradigm or framework at all, but rather as *freedom* from such an approach. In defining the so-called 'alternative paradigm stance,' Greene (2007) described the pragmatist view as *not* carrying underlying "philosophical assumptions", while Biesta (2010) describes pragmatism as a set of "philosophical tools" for addressing problems, rather than "a philosophical position," – a perspective which Tashakkori and Teddlie (2010) refer to as an "unparadigm". This freedom from rigid underlying perspectives is particularly beneficial to the multi-phase mixed methods approach employed in this thesis, allowing for an evolving and adaptive approach to the research process, rather than a prescribed notion of how the research should be conducted which may be at odds with later developments in the data gathering and analysis processes.

3.3 Experimental Design

The nature of Virtual Citizen Science and the form which sociality takes within VCS projects raises significant issues for quantitative experiments. On the whole, experiments within VCS tend to take one of two forms – A-B testing with small numbers of participants (see for example: Starr et al. (2014) and Sprinks et al. (2014)) or analyses of data gathered directly from Citizen Science projects in non-experimental conditions (see for example: Luczak-Roesch et al. (2014) and Tinati et al. (2015a)). Creating a Citizen Science project to host experiments was infeasible at the outset of this research and remains an extremely challenging prospect. VCS is distinct from other forms of crowdsourcing, in that participants are predominantly motivated by altruism and their own intrinsic interests in the underlying research topic, scientific domain, project task or project assets [4]. Any such initiative would likely fail to attract sufficient participation, as studies conducted within existing projects but from different scientific fields or with different research aims have previously been viewed negatively and rejected by project participants (Darch, 2017). Careful management of projects is also required to ensure the long-term viability of projects, as well as sufficient contributions of sufficient quality from the community (Curtis, 2018b). Even otherwise highly successful topics can be unpredictable and problematic for the purposes of extended research, such as in the case of the Andromeda Project, a highly effective Zooniverse project that ultimately lasted only a few weeks due

[4]For further details see chapter 4

to unprecedented interest from volunteers[5]. For this reason, in chapters 7 and 8, I draw upon data gathered from live VCS projects based on predetermined variables, rather than using a custom experiment or project.

3.4 Project Selection

As described in chapter 1, due to the decision to analyse data drawn from live projects rather than custom experiments, it became necessary to select the projects to be analysed. In spite of the diverse and growing selection of VCS projects, sourcing data from within VCS projects can be a challenge. Citizen Science is not necessarily open science and similarly VCS data are not open data – demonstrated by the fact that none of the projects identified in chapters 4 and 5 have made their data fully available to the public, even in the case of projects which have been retired and completed for many years. Where data *are* published, these are largely outdated by the time of release due to the lengthy publication cycle associated with VCS. Given the large datasets involved, it is also not feasible to analyse data from several projects.

Instead, I have selected two platforms to analyse, such that the results can be generalised to a large proportion of the VCS landscape. The first is the Zooniverse platform[6], which is the largest VCS platform[7] at the time of writing. The Zooniverse platform has also influenced and inspired a variety of subsequent citizen science projects in the domain of astrophysics and beyond, such as the Dutch National Archive Crowdsourcing initiative (Tovgaard-Olsen et al., 2019), VerbCorner (Hartshorne et al., 2013) and a citizen science initiative by the New York Public Library (Vershbow, 2013). Similarly, the Zooniverse has been viewed as a model example for the development of crowdsourcing technologies in the area of digital heritage (Hecker et al., 2018). The introduction of the Panoptes platform in 2015 has significantly streamlined the project design and publishing process and in doing so, has also partially homogenised the design and implementation of Zooniverse projects – in particular, the project builder restricts the availability of new features until they have been successfully trialled in existing Zooniverse projects (Home, 2019). The launch of this Zooniverse project builder has resulted in a significant rise in the number of Zooniverse

[5] https://blog.zooniverse.org/2013/12/05/andromeda-project-we-hardy-knew-ye/
[6] https://www.zooniverse.org/
[7] Both in terms of the number of projects (Spiers et al., 2019) and in terms of volunteer numbers (Muhtaseb, 2019).

projects, both within the platform and hosted elsewhere[8] (Rosser and Wiggins, 2019). By analysing data from the Zooniverse, I aim to not only analyse the largest share of VCS projects and volunteer numbers, but also to consider a large body of potential future projects, launched within the Zooniverse and hosted elsewhere, as well as those which may be influenced by the design of Zooniverse projects.

The second platform is EyeWire[9], which while smaller than Zooniverse boasts over 250,000 players[10]. The decision to draw on EyeWire was partially informed by the project review process, which showed that EyeWire shares many online community and gamification features that are present in some form in other VCS projects. Moreover, EyeWire has a much larger number of active players than other well known, gamified VCS projects – in FoldIt, for example, the core player base consists of around 200 users (Curtis, 2018b). To further aid in generalisability, the questions posed to the EyeWire team during the interview process were specifically chosen to establish the differences and similarities between EyeWire and other VCS projects and to gauge the team's views on how their experiences and findings may vary from other VCS projects – particularly gamified projects and Games With A Purpose. While the team's responses were ambiguous in terms of whether they had intentionally implemented features from other projects, their responses and the results of the project survey process all suggested a large degree of overlap between EyeWire and other gamified Games With a Purpose in terms of features and approach. Although the nature of the game task in EyeWire is relatively unique, the features which make it a game – badges, leaderboards, competitions and narratives – are all closely aligned with other Games With a Purpose and gamified citizen science initiatives.

3.5 Study Data and Methodologies - Qualitative Methods

Chapters 4, 5 and 6 of this thesis make use of a qualitative methodology drawing on a literature review, project review and interview approach respectively.

[8]Projects built through the project builder do not necessarily become 'official' or full Zooniverse projects and are not automatically hosted and indexed on the Zooniverse website

[9]`https://eyewire.org/`

[10]`https://www.irishtimes.com/news/science/gamers-helping-to-piece-together-the-brain-s-complex-circuitry-1.3452825`

This section outlines the data collection and analysis processes employed in each of these chapters. The research conducted within this thesis was approved by the ethics board of the Faculty of Physical Sciences and Engineering of the University of Southampton under ERGO number 25152.

3.5.1 Systematic Literature Review

Tashakkori and Teddlie (2010) recommend that the research process should begin with a literature review, as a means to ascertain and identify the "existing evidence" surrounding a given topic, problem or research area. In its simplest form, a literature review "means locating and summarising" different research publications and results – both quantitative and qualitative – with regard to a given subject (Cresswell, 2014; Tashakkori and Teddlie, 2010). There are numerous different approaches and methods which can be integrated in conducting a literature review and the process may be as simple as identifying desired keywords and conducting a search for these terms in whichever library or digital database a researcher chooses (Cresswell, 2014; Teddlie and Tashakkori, 2009). Such a method, however, is not without its weaknesses – particularly in the form of biases and errors. The method and criteria used by a researcher may introduce these biases or errors – missing publications or excluding particular findings due to the format of the data for example – but so, equally, may the publications that a researcher selects (Tashakkori and Teddlie, 2010; Webster and Watson, 2002). Some journals, for example, only publish secondary articles – for example literature reviews – and so any article that uses a literature review methodology with inherent biases will introduce these biases into subsequent reviews (Teddlie and Tashakkori, 2009). In turn, more prominent journals and commonly cited publications are likely to have a stronger and more significant contribution to the field and research question and should be given more weight during the data collection and synthesis processes than other, weaker articles (Webster and Watson, 2002).

The *systematic* literature review is a review method specifically intended to address and overcome these issues. A literature review is *systematic* if it makes use of a predetermined methodological approach for the selection, collection and assessment of literature sources (Tashakkori and Teddlie, 2010). Unlike other forms of literature review, a systematic literature review begins with an explicitly defined research question and search strategy, which is clearly documented, alongside specific criteria for the inclusion or exclusion of sources – a process which generally includes consideration of the quality and accuracy of the sources, as well as their content (Kitchenham, 2004; Tashakkori and Teddlie, 2010). On the other hand, in a systematic review of systematic reviews within the field of software engineering conducted by Kitchenham et al. (2009), just 3 out of 20 sampled publications offered quality assessment metrics or scores for samples sources and the majority provided no such information at all.

Databases and Search Terms

Although for the sake of simplicity this thesis uses the term "Virtual Citizen Science," there are in truth a variety of descriptors used within the literature and it is unclear which, if any, names dominate within the field. To avoid introducing sources of bias, searches were conducted for the three most commonly occurring terms within the background literature – that is, *digital*, *online* and *virtual* Citizen Science. A fourth term identified within some sources – *technology mediated* Citizen Science – was not included within the search process due to apparent ambiguity and a lack of clarity regarding the phenomenon and processes which it describes.

The primary criteria in selecting databases were the need for multidisciplinarity and the indexing of academic publications of high quality. Initially, the JSTOR, Scopus and Web of Science databases were selected based on these criteria and influenced by the recommendations and sources used by Seaborn and Fels (2015). A brief consideration of these initial results demonstrated two significant issues – firstly, that there was a high degree of overlap between the three databases and secondly, that while the databases were indeed multidisciplinary, a small number of disciplines dominated the results. The PubMed database was introduced to attempt to resolve this issue, but only identified a small number of additional publications. Google's Scholar database was then consulted in an effort to diversify the sources identified during the literature review process. The Scholar platform indexes a greater number of sources than the four previous databases, but features a much less robust search function and is not restricted to high quality research articles. Nevertheless, the sources identified through Scholar were more diverse than the other databases and the platform was therefore advantageous as a literature review tool.

Although the advanced search function was used for all five databases, the

Repository	Search terms	Results
JSTOR	((ab:("online" + "Citizen Science") OR ab:("digital" + "Citizen Science") OR ab:("virtual" + "Citizen Science"))	15
Scopus	TITLE-ABS-KEY ("online" + "Citizen Science") OR TITLE-ABS-KEY ("digital" + "Citizen Science") OR TITLE-ABS-KEY ("virtual" + "Citizen Science")	512
Web of Science	TOPIC:("online" + "Citizen Science") OR TOPIC:("digital" + "Citizen Science") OR Topic:("virtual" + "Citizen Science")	369
PubMed	(("digital Citizen Science"[Title/Abstract]) OR "Virtual Citizen Science"[Title/Abstract]) OR "online Citizen Science"[Title/Abstract]	9
Google Scholar	"online Citizen Science" OR "Virtual Citizen Science" OR "digital Citizen Science"	998

Table 2: Databases and search terms used and number of results identified through literature review process.

available options and enforced restrictions did not allow for completely identical search terms. Where possible, searches included the title, abstract and keywords attached to publications. In the case of PubMed and JSTOR, no keywords were available to search and the advanced function allowed only the title and abstract to be searched. It should be clarified that for JSTOR, as few as 10% of the publication records include abstracts, but no other suitable restrictions are available within the search interface and completing the search without any restriction returns far too many results to feasibly analyse fully. Web of Science does not feature such a restriction, but does feature the "topic" setting, which filtered papers in a similar and sufficiently accurate manner. Conversely, Scholar features no such filtering options and so more general search terms were used for this database. A full list of the specific search terms used for each database can be seen in table 2.

Inclusion and Exclusion Criteria

The following inclusion criteria were applied to selecting publications:

1. **English Language** – Publications written in a language other than English were excluded during the sampling process, as a means to prevent errors during analytical and synthetic procedures.

2. **Peer-Reviewed Sources**– As a partial proxy for quality, only those publications which had passed peer-review were included within the sampling process. This was predominantly confirmed by checking the journal or conference, or confirming the status of papers marked as arXiV preprints. Dissertations and PhD theses were excluded for similar reasons.

3. **Journal or Conference Papers** – Ascertaining the quality and validity of source publications is a difficult process. To assist with this process, publications were selected for inclusion only if they were originally published in conference proceedings or an academic journal. Workshop papers were rejected as a result of ambiguities in distinguishing between peer-reviewed submissions and non-peer-reviewed position or opinion papers. While in theory books based on peer-reviewed research would be of suitable quality, no sources were identified where relevance to VCS could be ensured.

4. **Empirical Research** – The aim of this literature was to identify empirical findings regarding the nature of VCS and engagement from the wider literature. For this reason, publications were included only if they contained original empirical research, or if they contained secondary synthesised findings based on original empirical research (such as a literature review or meta synthesis)

5. **Virtual Citizen Science** – Publications were included only if they made clear reference to Virtual Citizen Science, with results from VCS projects drawing on active participation from volunteers. Publications which could not be confirmed to draw on data or experimentation involving VCS projects, or which used VCS only as an example along other forms of crowdsourcing were removed from the sample.

6. **VCS Project Introductions and Results** – A large proportion of publications presented only the results of individual VCS projects, such as galaxy catalogues from Galaxy Zoo. These publications used domain-specific concepts, terminology and methodologies which require specialised training in a given field (e.g., astrophysics) to understand and which I would not be able to accurately analyse and assess. Moreover, some project introductions offered little in the way of a contribution to the findings

and arguments put forward during the literature review process. For these reasons, these were removed from the sample when identified.

7. **Applications of VCS Projects** – VCS projects have been applied in a number of areas, such as education and training in classroom or employment settings. Publications describing these applications predominantly assessed the effectiveness of the application process, rather than evaluating or assessing the VCS projects themselves and so were outside of the scope of this thesis. Furthermore, in many cases the VCS project could be viewed as interchangeable with other methods or materials and little in the results distinguished the impact of VCS from the impact of other methods. In addition, since projects were most commonly applied to education and training, it was not possible to distinguish projects which represent *Virtual* Citizen Science from projects which would better fit the *education* category set out by Wiggins and Crowston (2011).

8. **Applications of VCS Results** – A large proportion of the sampled literature contained projects detailing the usage of data from VCS projects for additional purposes such as the training of machine learning algorithms. While these applications represent VCS-based research such as training algorithms to detect galaxies based on Galaxy Zoo data, as with other applications-based papers, these offered little in the way of information about the VCS data or processes themselves and were therefore also excluded from the sample.

9. **Mature Results** – Publications had to include mature results, with evidence-based conclusions and a discussion of how these results fit within related literature and the wider field. Publications with the word "towards" in the title or which did not have full results were eliminated during the second round sampling process.

Data Collection Process

As a means to conduct the sampling process, a three stage procedure was used. During the first round, a search was conducted of each of the five databases using the search terms set out in table 2 and a full list of publications extracted. As there was a strong degree of overlap between the five databases, this initial process also aimed to identify repeated articles within the sample by identifying identical titles, abstracts and metadata. A total of 450 of the articles were deemed to be duplicates and these were removed from the samples to prepare for stage two.

Upon the extraction of the majority of the duplicated publications, a second and more detailed process began. The aim of this stage was to filter the large number of publications into a more manageable and smaller sample, as efficiently yet accurately as possible. To this end, each title, abstract and accompanying metadata – such as keywords – were examined and papers were assessed for suitability based on the exclusion criteria detailed previously. A full list of the different categories to which publications were assigned can be seen in table 3, while the papers removed during round 1 can be found in appendix A, grouped according to the reason for their removal in table 3. Where these fields alone were insufficient to identify whether a paper should be excluded, the paper was included so that more detailed analysis could be carried out during the third stage of the collection process. The one exception to this concerned publications which could not be confirmed as peer reviewed. Google searches were carried out to confirm the nature of the journals, conferences and other source materials and if a paper could still not be confirmed as peer reviewed, it was removed from the sample at this point.

Finally, during the third stage, each paper was read in full multiple times. During the first reading, the paper was further assessed for relevance and suitability with regard to the inclusion and exclusion criteria. Table 4 describes the full list of reasons for the removal of publications during round 2, while each of the specific papers is detailed in appendix B, catalogued according to the reasons for removal. After this stage, a second reading was used to code the paper using the NVivo software package, with nodes detailing the methodology used, projects mentioned, main findings and outcomes and any other significant outcomes or details of note. A third and final reading was carried out at the end of the coding process for each coded paper, to ensure consistency and accuracy.

The initial stage was not intended to identify all duplicates, particularly as some papers were listed multiple times but with different titles or author listings. Instead, steps were taken during stage two and three of the sampling process to identify these sources. In some cases, publications were not exact duplicates, but consisted of a journal article developed from a previous conference paper, or two journal papers of similar quality. Where this occurred, the more detailed of the two publications was selected for the sample and the less detailed was excluded – generally speaking, but not always, this meant the exclusion of the earlier paper. A summary of the final results of the literature review process can be found in chapter 4 in table 5.

3.5.2 Project Survey

As previously outlined, it would be impossible to examine every possible VCS project. Early projects have been retired, relaunched or replaced (see for example, the list published by Tinati et al. (2015b)). Furthermore, in the time that the research described in this thesis was conducted, there has been a rapid rise in the number of VCS projects within the Zooniverse platform due to the launch of the *Panoptes* service, which allows any interested stakeholder to design, launch and manage their own VCS project (Bowyer et al., 2015; Kosmala et al., 2016). Identifying new and existing projects is also difficult, as the vast majority are not the subject of published research materials and there is no central database listing projects or platforms[11] (Kullenberg and Kasperowski, 2016).

During the literature review process described in chapter 4, all listed projects were extracted from the final sampled literature and these can be found in table 6 in chapter 4. Unfortunately, the number of projects identified in this way was small and the diversity was relatively low, with a small number of projects and platforms appearing multiple times. Moreover, many of these projects were at a late stage in their lifecycle or had already retired. In the intervening months since the project review was conducted, many of these projects have been relaunched in a different form, but at the time of conducting the project review, the list was generally unsuitable for project review purposes.

Instead and as a means to identify more diverse and unpublished projects, I followed the advice within the literature that participants can suitably spread news of projects through word-of-mouth and turned to citizen-led repositories (Darch, 2017; Kraut et al., 2012). Tinati et al. (2015b) explicitly mention the spread of projects and assets through social media, but unfortunately finding such comments is difficult without knowledge of the project name to search for or access to groups of VCS volunteers. Ultimately, the most well maintained list

[11] While there are services such as Sci Starter and organisations such as the Citizen Science Association, there is no requirement or guarantee that projects will be registered with such a service and Sci Starter in particular has a relatively simplistic search function that makes identifying projects sharing specific characteristics difficult.

identified was Wikipedia's List of Citizen Science Projects [12]. Naturally, since Wikipedia is a volunteer-led online encyclopaedia, anyone can add information to the site and all projects were carefully analysed by visiting the project URL, where listed, as well as through Google and Google Scholar searches to seek further details on projects.

Online Community Framework

To identify the extent to which projects adhered to online community design recommendations, it was first necessary to synthesise a framework of recommendations based on empirical evidence from existing literature. The aim of this process was not to identify an exhaustive list of all possible recommendations, but rather to identify key themes and related affordances for the successful design and implementation of online communities and community-based features and tools. A pilot literature review was initially conducted using the same databases in table 2, using the search term "online" + "communit" followed by the database specific wildcard character to identify literature covering online community and online communities – e.g., "online" AND "communit*". However, it became clear that a second literature review would be a time consuming and inefficient process, due to the sheer volume of online communities literature available and the fact that many of the sources contained recommendations and observations that were specific to one or more types of community.

As a solution to this, the decision was made to use a secondary, review-based source to identify recommendations from a wide-range of sources efficiently. For this purpose, I drew on the publication *Building Successful Online Communities* by Kraut et al. (2012), which presents qualitative and quantitative empirical evidence from a variety of academic publications, across a diverse set of online communities – discussion forums, crowdsourcing and Citizen Science platforms, MMORPGs such as *World of Warcraft* and even retail services such as *eBay*. The publication is particularly suitable for use during this process as it separates recommendations into five key categories[13], allowing for the extraction of only those recommendations which are of relevance to questions of productivity and engagement.

[12] `https://en.wikipedia.org/wiki/List_of_citizen_science_projects`

[13] Encouraging contributions and productivity, encouraging long-term commitment, managing and regulating behavoiur, attracting and managing new participants and launching a new community

Additionally, three frameworks of community success were employed, to offer further perspectives or fill gaps within the findings presented by Kraut et al.. The first of these was the model outlined by Preece (2001), which similarly outlines key factors which influence the likelihood of participants contributing to a given project. Iriberri and Leroy (2009) approach this issue in a different way, instead offering symptoms of unhealthy or ineffective communities and a lack of participant contribution. Finally, the work of Nov et al. (2014) as identified through the literature review process was included in the theme identification process as the only source that had explicitly drawn links and conclusions between specific motivational factors and contribution outcomes.

I first extracted or synthesised design recommendations from the literature. In the case of the recommendations made by Kraut et al. (2012), this process was relatively simple, as the publication contains explicit design claims - for example, "fear campaigns cause people to evaluate the quality of persuasive appeals". For each recommendation, the main focus of the recommendation, the outcome and the nature of the recommendation were summarised – in this particular case, the main focus would be fear campaigns, the outcome is that participants will question the appeal and the nature of the appeal is negative. Recommendations were then further compared with the findings of Iriberri and Leroy (2009), Nov et al. (2014) and Preece (2001), with divergent recommendations extracted and added to the list of observations and themes. A full list of summarised recommendations can be seen in appendix C.

To ensure relevance to success in Citizen Science, I selected those recommendations with a focus on ensuring a high quantity and quality of contributions. This was in keeping with the most commonly occurring success measures, as well as concepts identified through the literature review process, while also being a strong focus of the underlying online communities literature used within the framework. Further, I selected only recommendations observed from the systems alone; recommendations which would require consultation with participants were deemed to be too complex due to the number of projects involved, but also due to the desire to explore the key concept of design, rather than user experience, which my success framework suggests is influenced by a range of factors beyond the design process.

- *Task visibility* – the ease with which participants can see and select microtasks and discussions requiring completion, but also share completed microtasks and contributions with others. A key component of ensuring the quantity of contributions and ensuring that required work is completed.

- *Goals* – the provision of challenges and targets for participants to achieve. Predominantly related to the quantity of contributions and number of users involved.
- *Feedback* – mechanisms for informing participants of the quantity or quality of submissions. Predominantly pertains to informing volunteers of the quality of contributions, although the framework also identifies the possibility of feedback related to the number of contributions/volunteers involved with a project.
- *Rewards* – tangible or intangible awards given to participants for making contributions or achieving goals. May be based on the quality of contributions (performance-based) or on the number of contributions made, although in rare cases may also be based on the number of volunteers engaged with a system.

Structured Walkthroughs

Ideally, the project review purpose would make use of a virtual or trace ethnographic approach, observing participants interactions within the virtual environment, as used by Curtis (2015b) in the case of FoldIt. Unfortunately, such an approach is incompatible with the majority of the sampled projects, which do not allow for tracking and observation of volunteer submissions. Instead, to assess these themes within each of the selected projects, I simulated this process, conducting the survey by utilising a structured walkthrough-based approach. For each project, I registered as a participant and completed approximately ten classifications (or for data collection projects, assessed existing contributions), as well as observing community interactions. This was a four step process: Initially, I registered and completed 10 classifications or other forms of contribution (evaluating existing contributions in data collection projects) within each of the 48 projects analysed within the study. During this initial step, I produced a list of affordances, mechanisms and characteristics observed across the projects. Following this, I examined points of contention to produce a common, unified list of mechanisms. I then returned to each of the projects, completing further contributions where necessary, in order to survey the number of times each of the mechanisms within the list was utilised within the sampled projects. Upon completion of this survey, I again corrected errors and identified disagreements, to produce a final list detailing the observed mechanisms and the number of occurrences of the mechanism across the projects. In order to ensure accuracy and to prevent possible issues with the structured walkthrough approach, this list was then compared with evidence drawn from the literature review process, using relevant publications where available for each project, as well as project blogs and news feeds.

3.5.3 Interview

Throughout the interview process, I followed phenomenological principles as a means to inform the research design, data gathering and data analysis process. Phenomenology is a philosophical perspective based on the study and interpretation of individuals' potentially subjective *Lebenswelt* – 'life worlds' or conscious experiences – which are highly influenced by the language context and culture of the society in which that individual finds him- or herself (Tashakkori and Teddlie, 2010). Phenomenological research describes predominantly qualitative, interview-driven research during which a researcher elicits the lived experiences of members of a particular community who have all experienced a given common phenomenon (Cresswell, 2014). By combining perspectives from several individuals, the researcher builds up a descriptive image of the phenomenon being studied, based on his/her intuition, reflections and judgements of the subjective facets described by the interview participants (Teddlie and Tashakkori, 2009). At the same time, the number of participants assessed through phenomenological studies is small and the subject of the study is highly focused – Cresswell (2014) advises between 3 to 10 individuals. Due to the subjectivity of phenomenology, it provides insights and experiences only on the specific phenomenon as experienced by the research participants in a specific context (Cresswell, 2014). For this reason, the interviews within this chapter offer distinct experiences of one individual project – EyeWire – based on participants' roles and activities within the project, to ensure that the phenomenon is common to – and understood in roughly similar terms by – each participant. As Converse (2012) describes: "The goal of phenomenological research is not to create results that can be generalised, but to understand the meaning of an experience of a phenomenon."

Data Collection

The phenomenological research paradigm generally requires semi-structured or unstructured interview formats (Ray, 1985). For this reason, each interview made use of a relaxed, semi-structured framework. In addition to a predefined set of questions, participants were enabled and encouraged to discuss their experiences at length, departing from the subject of questions where necessary. This structure was similar to – and inspired by – the method described by Tinati et al. (2015b).

Three semi-structured interview sessions were held with members of the EyeWire team during February 2017. In addition to an interview schedule consisting of 24 questions, participants were encouraged to contribute their experiences, recommendations and opinions, both in response to these questions and to external topics. The full interview schedule can be found in appendix E. Interviews were carried out on site at the EyeWire offices in Boston, Massachusetts in the United States across two working days in late February, 2017. On the first day, an informal and unrecorded group session was held during which the aims of the research, as well as any ethical concerns were discussed with the group. The predominant aim of this session was to assist participants in remembering the potentially lengthy history of EyeWire, to develop a common understanding of the terminology and concepts associated with EyeWire and also to ensure that participants would be able to access any necessary materials to aid in evoking and explaining their experiences during the sessions themselves.

Six members of the EyeWire team took part, across 4 roles, which represented the entirety of the full-time EyeWire staff at the time at which the interviews took place – with the exception of one full-time developer who was unavailable for interview. EyeWire staff work in three distinct teams, situated in one of three offices with distinct job roles sharing high amounts of intra-office synergy. For example, the three game masters have specific and distinct tasks within the project, but work closely together and are highly reliant on one another to successfully complete project tasks. To take full advantage of these synergies and to ensure greater accuracy and capacity for remembering experiences, each interview session was conducted with these smaller office-based teams. As a result, each interview session varied in terms of the number of participants:

1. Community and Project Manager – Responsible for overseeing the project

and the team, as well as external partnerships and dissemination – This participant was interviewed alone during interview session one, due to her wealth of knowledge and greater experience of the project and the nature of the role as a cross-sectional member of both sub-teams and independence from the team structure.

2. Designer – Responsible for designing artistic features, interface elements and the project 'brand' as a whole – This participant was interviewed alongside the developer during interview session two.

3. Developer – Responsible for implementing new game, hardware and software elements, resolving system bugs and maintaining the game – This participant was interviewed during session two alongside the designer.

4. Game Master – Responsible for interacting with the community, planning and running community-engagement activities and eliciting feedback from the community. Interviewed during session three.

5. Game Master – Interviewed during session three.

6. Game Master – Interviewed during session three.

Each interview was recorded with permission from participants and recordings were later transcribed for analysis, in line with recommendations for interview-based research and particularly phenomenological research methods (Cresswell and Plano Clark, 2018).

Data Analysis

Upon transcription of the interview findings, analytical processes followed a standard workflow for phenomenological studies (Goulding, 2005; Wojnar and Swanson, 2007). This workflow, first codified by Collaizzi (1978) is an iterative, 7-stage process which involves the following processes:

1. The researcher reads the interview transcripts numerous times.

2. The researcher identifies key phrases, terminology and statements from the transcripts which pertain to the phenomenon under study.

3. The researcher assigns meaning to the significant terms and statements identified in stage 2.

4. The researcher repeats this process for each participant's experiences, clustering common statements and experiences where appropriate.

5. The researcher 'integrates' these clusters, commonalities and statements into a coherent description of the phenomenon under study.

6. The researcher reduces these themes into a description of behaviour or activity.

7. If necessary, the researcher returns to the research participants to clarify or request feedback on the codified themes and descriptions.

Each transcript was coded using NVivo, using an iterative, multi-stage process. After reading each transcript through multiple times, key phrases and statements were highlighted for each transcript and compared to ensure consistency. After this, the phrases were clustered into common themes and meanings. Since these interview findings built upon observations from the literature and project review processes, these initial themes aligned with themes and areas from chapters 4 and 5 (see below). After clustering the phrases into themes, each transcript was read once again and the highlighting process was repeated, adjusting phrases which no longer fit within the clusters and themes and adding phrases which fit within the clusters and had not previously been added. The now finalised clusters were then integrated according to commonalities not only in the clustering, but in the statements and passages themselves and written up into summaries of the behaviours and decision making activities exhibited by participants and the design team.

In accordance with recommendations from the wider literature[14] draft versions of the descriptions and analyses generated from the interviews were sent to the community manager to request any corrections and highlight any invalid interpretations of these analyses. No such comments were received from the team. To ensure validity, where possible, the descriptions and analyses contained within this chapter compare the findings, descriptions and examples offered by the team with findings from the literature review conducted within chapter 4.

The findings are presented in chapter 6 grouped according to the final clusters and themes, which are:

[14] see for example: Wojnar and Swanson (2007)

- Project Needs and Success Criteria
  - Efficiency and Cost-Effectiveness
  - Fun and Engagement
  - Public Dissemination and Awareness
- Participant Motivations
  - Altruism
  - Intrinsic Factors
  - Social and Community Factors
  - Extrinsic Factors and Rewards
- Online Community Features
  - Task Visibility
  - Goals
  - Feedback
  - Rewards
- Role of Sociality
  - Game Management
  - Enabling Tasks
  - Relationship with Productivity

3.6 Study Data and Methodologies - Quantitative Methods

In contrast to the qualitative methodologies detailed above, chapters 7 and 8 of this thesis draw on quantitative methodologies and concepts, addressing outstanding questions raised by the results of the initial chapters. These include hypothesis testing and the calculation of effect sizes, including correlation.

3.6.1 Hypothesis Testing

Hypothesis tests employ a *null hypothesis* – a hypothesis in the form of a null statement, that factor f has no significant impact on population μ (Rouder et al., 2009). Outcomes are then judged based on statistical trends, often utilising the p-value, which represents the probability that a given result would occur by chance in a given population. For example, a p-value of 0.05 represents a 5% chance of a given result being seen in a population, or a 5% false positive rate (Duncan, 1975).

The t-test is the most common test of the statistical significance between the means of two samples or populations (Fagerland, 2012). The test outcome – the t-statistic – corresponds to a given p-value based on the normal distribution (bell curve), such that any t-statistic can be associated with the likelihood that a result would be observed by chance and the null hypothesis accepted or rejected on this basis (Rouder et al., 2009). It is a *parametric* test, which means it carries certain assumptions about the underlying data that must be met for the test results to be accurate. In the case of the t-test, the statistical assumptions are that the data are derived from a normally distributed sample or population and the robustness or suitability of the test for non-normally distributed populations – which represent the majority of cases – is a subject of much research and debate (Johnson, 1978). Moreover, the t-test assumes equal variances within both samples or populations, although there is a commonly used alternative known as "Welch's Approximation" for samples with unequal variances (McElduff et al., 2010). The formula for the t-test is (Johnson, 1978):

$$t = \frac{\bar{X}_1 - \bar{X}_2}{s \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \tag{3.1}$$

Where $\bar{X}_1$ and $\bar{X}_2$ represent the arithmetic mean of two samples 1 and 2, n_1 and n_2 represent the sample size of samples 1 and 2 and s represents the standard deviation of both samples.

Welch's equivalent test for unequal variances is given by (Delacre et al., 2017):

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\frac{s}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}} \tag{3.2}$$

Where $\bar{X}_1$ and $\bar{X}_2$ represent the arithmetic mean of two samples 1 and 2, n_1

and n_2 represent the sample size of samples 1 and 2, s represents the standard deviation of both samples and s_1^2 and s_2^2 represent the variance of samples 1 and 2 respectively.

While the t-test is arguably robust for larger sample sizes due to the central limit theorem[15], the test is nonetheless arguably less robust than non-parametric tests for non-normally distributed samples (Duncan, 1975; Norman, 2010). There are various non-parametric tests, which are designed and intended for samples with specific characteristics, such as the sign test, based on the median of a population, the Wilcoxon Signed Rank Test for the median of the differences between paired, measurement-based data and the Mann Whitney-U test for the median of differences between unpaired samples (Rees, 1989). In this thesis, the sample data are unpaired and do not meet the assumptions of the t-test and so the Mann Whitney-U test is the most suitable hypothesis test for this purpose.

The Mann-Whitney U test, generally seen as the nonparametric equivalent of the t-test, is a test for whether two sampled observations were taken from the same population and share the same distribution (MacFarland and Yates, 2016; Mann and Whitney, 1947). While not as strong as the t-test, the Mann Whitney-U test closely approximates the t-test in terms of accuracy, with approximately 95% of the accuracy of the parametric equivalent, making it one of the most powerful non-parametric tests (Landers, 1981). For this reason, the Mann-Whitney U test is best suited for the calculations used in this thesis, which have variable sample sizes and do not meet the conditions of the t-test. The formula for the test is given by (Zar, 1996):

$$U_1 = R_1 - \frac{n_1(n_1+1)}{2} \tag{3.3}$$

Where U_1 represents the U statistic output by the test, R_1 represents the sum of the ranks for sample 1 and n_1 represents the sample size of sample 1.

3.6.2 Effect Size

Effect sizes are “the most important outcome of empirical studies,” providing standardised measures of the impact that a given factor has on a sample or population, allowing comparisons between the findings of separate studies and

[15]The theorem that as the size of a population (and associated samples) increases, so the distribution of the sample approaches the normal distribution (Rosenblatt, 1956)

also providing vital details to inform subsequent studies within a research area (Lakens, 2013). In fact, Cohen (1990) suggests that the most significant aim of hypothesis testing research should not be to report or prove statistical significance (i.e., the p value) at all, but rather to report effect sizes. There are numerous methods for calculating, identifying and reporting effect sizes for statistical tests, many of which are associated with specific processes or situations. For example, η^2 is an effect size most commonly associated with the ANOVA statistical test (Bakeman, 2005). The purpose of this section is not to outline each possible measure of effect size, but rather to identify the most significant measures for this research and to explain and justify their use.

Cohen's d is a measure of effect size for comparing two conditions, such as in the case of the t-test (Fritz et al., 2012). It is a parametric statistic, drawing on the mean and standard deviation of a population to identify effect sizes. Cohen and later Sawilowsky (2009) have provided "rules of thumb" that allow for the conversion of numerical effect sizes into verbal indicators of small, medium and large effects. The main advantage of the use of d is that it is a commonly used effect size and is often used to express the results of t-tests due to meeting the same underlying statistical assumptions (Lakens, 2013; Rice and Harris, 2005). As a result, it allows for easy comparison between results, particularly in the field of psychology where d finds common use (Lakens, 2013).

r is a non-parametric measure of effect size, recommended by Cohen (1988) for use with non-parametric tests such as the Mann Whitney-U test. Rather than using the standard U statistic, it is based on the z statistic which may also be calculated and reported using the Mann Whitney-U test[16]. It is calculated using the following equation:

$$r = \frac{z}{\sqrt{N}} \tag{3.4}$$

Where z is equal to the z-statistic output by the Mann Whitney-U test (or other nonparametric test) and N is equal to the combined sample size of the two samples used to calculate z (Fritz et al., 2012). In turn, z can be calculated

[16]By converting U to z, it is possible to identify the statistical significance p as an outcome of the test in the same way one could use U. The calculation of z is particularly suited to comparing two samples where N>8 (Nachar et al., 2008).

using the following equation:

$$z = \frac{U - \frac{n_1 n_2}{2}}{\sigma U} \tag{3.5}$$

Where U is the U statistic, n_1 and n_2 represent the number of observations in samples 1 and 2 respectively and σU represents the standard deviation (Nachar et al., 2008).

Although direct calculation of *d* is unsuitable for nonparametric tests, as it assumes normality, *d* can be easily calculated from nonparametric equivalent measures of effect size (Fritz et al., 2012). For example, by calculating the z score, one can then easily infer the measure *r* from which *d* can then be calculated. While this is somewhat more work than simply calculating *r*, it is desirable for this thesis as in revisiting and standardising *d*, Sawilowsky (2009) provides 5 categories (very small, small, medium, large and very large respectively), while only 3 categories were specified by Cohen for *r* and even a relatively small value of just 0.1 is sufficient to be considered a 'small' effect size (Coolican, 2017). It is therefore somewhat simpler to distinguish between otherwise similar effect sizes with *d* than with *r* – particularly larger and smaller effects.

3.6.3 Correlation Coefficients

Correlation examines the degree of a relationship between two or more variables, with at least one *dependent* variable and at least one *independent* variable (Tashakkori and Teddlie, 2010; Teddlie and Tashakkori, 2009). For the sake of this thesis, the type of correlation employed will be linear correlation, assuming a linear positive or negative relationship between a single dependent and single independent variable. This decision is supported by the use of scatter plots, which demonstrate a broadly linear relationship between the dependent variable of task contributions and the independent variable of talk contributions on any given day. There are three commonly employed linear correlation coefficients, which are described below. While the method for calculating each correlation coefficient varies, the resulting value in all cases represents a value on a scale between -1 and 1, where -1 represents a "perfect inverse relationship" between the dependent and independent variable, while 1 represents a "perfect direct relationship" (Mukaka, 2012). A value of 0 indicates that there is no linear relationship between the variables, although even a value other than 0 does not necessarily indicate a significant linear relationship between the two variables, particularly where the sample size is low (see section 3.6.3 below) (Taylor, 1990).

Pearson's r[17] is arguably the most commonly used correlation coefficient statistic within statistics. Pearson's r is a parametric statistic with a calculation that relies on two sample means and as a result, generally assumes that the source data points follow a normal distribution (Alexopoulos, 2010). As a result, the measure is extremely sensitive to outliers and even small amounts of contamination – as little as 10% and noise within the source data can quickly lead to high levels of bias in the resulting calculations (Abdullah, 1990).

$$r = \frac{\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \overline{x})^2(y_i - \overline{y})^2}} \tag{3.6}$$

Spearman's ρ is a non-parametric measure of correlation between two values. It is similar to Pearson's r, but is distinct in that it uses a non-parametric transformation of the two values according to a given function (Fredricks and Nelsen, 2007). As a result, Spearman's ρ is equal to r, but between ranks rather than between untransformed values (Rupinski and Dunlap, 1996). Spearman's ρ is less sensitive to outliers and extreme data points than Pearson's r and is more accurate when the sampled data points are non-normally distributed (Mukaka, 2012):

$$\rho = 1 - \frac{6\sum d_i^2}{n(n^2 - 1)} \tag{3.7}$$

Kendall's τ, as with Spearman's ρ is a non-parametric measure of correlation between two paired values (Long and Cliff, 1997). It represents "the probability of observing concordant and discordant pairs...rescaled from -1 to 1," in the same format as other measures of correlation (Lapata, 2006). τ is less commonly reported than Spearman's ρ, but has a number of significant advantages. Firstly, it is significantly less sensitive to outliers than ρ and r (Abdullah, 1990). Moreover, τ is less biased than ρ when the sample or population size is particularly small (less than 30) or large and is less prone to false negative rates – rejecting the null hypothesis when it is in fact true – than ρ, in line with the 95% confidence interval (Lapata, 2006; Xu et al., 2013). In turn, the values of τ reported within the literature are generally lower than those reported for ρ and

[17]Not to be confused with r as reported previously, which is a distinct measure of effect size with a similar name but different method of calculation and application.

r^2.

$$\tau = \frac{2S}{n(n-1)} \tag{3.8}$$

Interpreting Correlation Coefficient Values

Correlation coefficients can be seen as a second form of effect size measure that is highly dependent on the sampling method and sample used within the underlying calculation (Cohen, 1988). Moreover, while all correlation coefficients give a value between -1 and 1, the exact calculation and interpretation of this figure is variable and the outcome from initially similar calculations, such as ρ and τ can vary greatly based on the sample size and to a lesser extent, distribution (Lapata, 2006). Fundamental differences in the nature of these measures therefore make conversions between different correlation coefficients undesirable and inaccurate (Walker, 2003). Any interpretation of a correlation coefficient is therefore highly dependent on the coefficient used, the sample size and method and the context in which the coefficient is employed and where specific definitions have been applied to these effect sizes, they are highly disparate – consider, for example, the 0.14 and 0.5 defined by (Walker, 2003) and (Cohen, 1988) respectively as measures of a "large" effect size. Interpreting correlation coefficients is therefore a complex process and a one-size-fits-all approach is likely to lead to interpretation errors.

One solution to this issue is the use of the coefficient of determination r^2, which is equal to the square of Pearson's r. This represents the percentage of the variation within the dependent variable that results from variation in the independent variable and is therefore a somewhat better measure of causation than the simple correlation coefficient (Taylor, 1990). This measure has a number of disadvantages, however. Firstly, it is based on linear regression equations, which assumes a normal distribution of residuals[18] and homoscedasticity of data points, which are themselves similar assumptions to Pearson's r (Jarque and Bera, 1980). There are no equivalent values for Spearman's ρ or Kendall's τ and indeed such calculations should not be used as they provide a skewed coefficient of determination value (Murray, 2013). While calculations have been put forth to allow the conversion of ρ and τ into r, the underlying assumptions of r and

[18] Residuals can be considered as errors or unexplained measurements within the data. A residual is the difference between an observed measurement and the predicted value of that measurement according to the regression model (Darlington and Hayes, 2016).

r^2 will nonetheless not be met by any such value (Walker, 2003). Moreover, r^2 is rarely reported and shares the same issue as the coefficients, in that there is no commonly defined method for interpreting the results. Certainly for the purposes of this thesis, the coefficient of determination is unsuitable, as the assumptions concerning normality and homoscedasticity associated with r, r^2 and linear regression cannot be met.

To resolve these issues, for the purposes of this thesis I will use the following categories to describe the values of the coefficient value, as recommended by Mukaka (2012):

- Very high correlation - A correlation coefficient of between 0.9 and 1 or -0.9 and -1.
- High correlation - A correlation coefficient of between 0.7 and 0.9 or -0.7 and -0.9
- Moderate correlation - A correlation coefficient of between 0.5 and 0.7 or -0.5 and -0.7
- Low correlation - A correlation coefficient of between 0.3 and 0.5 or -0.3 and -0.5
- Negligible correlation - A correlation coefficient of between -0.3 and 0.3

It is, of course, important to underline that correlation does not and must not be taken to imply causation. Spurious correlation is a very real issue, such as a somewhat famous example of a given geographical area in which the number of storks closely correlates with the number of babies born (Grichting, 1992). The decision to calculate correlation between daily task and talk submissions within this thesis is based on evidence and findings drawn from the literature review process described in chapter 4 and the assumption is made on these grounds that any observed correlation is not spurious. Nevertheless, a causative relationship between the two factors is not directly assumed, but rather that any factors that lead to increased levels of task productivity will also drive talk activity.

3.6.4 Data Extraction

For the hypothesis testing process described in chapter 7, two separate datasets were used. The first, extracted directly from the EyeWire API, consisted of quantitative values for daily cube and point scores for every volunteer who had signed into the EyeWire system within the sampled period, which covered August 2017 to August 2018, as well as values for each formally recorded competition which had taken place since 2014[19]. The specific dates sampled were selected to avoid periods during which the team had offered prizes for participation, as while the team had not observed a negative effect after removing prizes, it is likely that there would be a significant effect on participant engagement. Each day was classified according to one of four categories – days on which minor competitions were held, days on which major competitions were held, days on which other kinds of competition were held (e.g,. the Happy Hour) and days on which no competitions were held. Since the EyeWire API does not provide timestamps for the submission of individual cubes, more specific information could not be ascertained regarding when a cube was submitted. Competitions in EyeWire broadly run from 9am EST on one day to 9am EST on another day. However, there are factors which encourage participation on competition days outside of the competition periods. The overall scoring system runs from approximately midnight to midnight and competition point rewards are counted in participants' daily scores; participants can earn bonuses from cubes completed before the competition commenced if these cubes are scythed or completed by other players and in the case of the marathon competition in particular, scoring is somewhat inexact and participants can often continue to earn bonuses after the marathon has formally closed[20]. As a result, the hypotheses and analyses conducted within chapter 7 consider competition *days*, rather than the specific competitions themselves, as separating out the specific competition score within EyeWire from work consciously done by participants outside of competition times is not possible.

The project data for the high-pressure projects detailed in chapter 8 were received directly from a member of the Zooniverse management team. These consisted of four database files, with the first two files covering task submissions and talk comments respectively made between the 28th of March and the 3rd of April 2017, while the second files covered task submissions and talk comments

[19]This represents the date when competitions were formally integrated into the EyeWire code-base. Before this point, competition infrastructure was artificial and relied on external tools and technologies.

[20]The marathon competition class is sometimes extended beyond 24 hours and the scoring, unlike other competition classes, is not performed live as the scores shown in the integrated leaderboard are manually updated periodically and are therefore not necessarily reflective of a participants' true score.

respectively between the 3rd of April and the 7th of April, 2017. Each file was first filtered using the software program Open Refine, to remove the large number of irrelevant data points from each file, such that only the submissions from each of the two desired projects remained.

Both task submission files contained the following fields:

- ID - An individual ID for each classification
- Source - Since all projects now use the Panoptes platform within Zooniverse, each source was identical and this field was removed
- Created_At - A timestamp detailing the date and time at which each classification was made
- Country_Code - The country code for the location from which the submission was made
- Country_Name - The name of the country covering the location from which the submission was made
- Project_ID - The four-digit code of the project to which the submission was made
- Subject_ID - The eight-digit code detailing the asset to which the submission corresponds
- Subject_URL - The URL of the asset to which the submission corresponds
- User_ID - A pseudonymous unique code assigned to each user, which could not be connected to an individual user's username or personal information
- Workflow_ID - This field overlapped in all cases with the Project_ID field and was also removed

The talk comment files contained the following fields:

- ID - An individual ID for each comment
- Source - As with the task submission files, this simply identified that the submission was made through the Panoptes system and this field was deleted

- Board_ID - The sub-board (e.g., science/help/chat) within which a comment was made. Each project has at least 4 sub-boards.
- Body - The content of the comment
- Body_HTML - The content of the comment in HTML format
- Created_At - A timestamp detailing the date and time at which each comment was made
- Discussion_ID - An individual numerical identifier attached to each thread, such that a parent comment and all replies share a single Discussion_ID
- Focus_ID - Identical to Subject_ID above. Only given where Focus_Type is not null.
- Focus_Type - Either *subject* or null, depending on whether the participant is commenting with regard to a specific asset or not.
- Country_Code - The country code for the location from which the submission was made
- Country_Name - The name of the country covering the location from which the submission was made
- Project_ID - The four digit code of the project to which the submission was made
- Section - Indicates whether a comment was made through a project or through the central Zooniverse talk page.
- Subject_Created_At - A timestamp showing when the asset was uploaded to the Panoptes system
- Subject_Image - URL of image, but distinct from Subject_URL above
- URL - The URl of the specific comment within the Zooniverse site
- User_ID - As above

3.7 Summary and Conclusions

This chapter has set out the philosophical and methodological concepts and principles to be followed throughout this thesis: mixed methods and pragmatism, both in the context of philosophy and of pragmatist research. I have described both the qualitative and quantitative methods to be used in each of the five following chapters, as well as the associated data collection and analysis processes. The following chapter draws on the qualitative literature review method to present the first of the five contributions of this thesis.

Count	Category	Description
70	Applications	Publications describing applications of VCS projects or activities
4	Crowdsourced Science	Publications describing crowdsourced science broadly
42	Duplicates	Repeated publications
43	Examples	Publications using VCS as an example of an idea or concept
10	Fringe VCS Topics	Publications describing irrelevant VCS topics (e.g., public dissemination)
188	Irrelevant	Publications with no relevance and reference to VCS
71	Methodology	VCS as methodology only
3	My Publications	Publications where I am an author
39	Not English	Publications written in a language other than English
51	Not Peer Reviewed	Publications which could not be confirmed as having been peer reviewed
62	Not Research Articles	Publications which were not scientific research articles (e.g., news, reviews)
118	Offline	Publications focussed on offline forms of Citizen Science
130	Other Crowdsourcing	Publications discussing other forms of crowdsourcing
11	Other Related	Publications discussing other, similar activities (e.g., open science, VGI)
99	Results	Publications only presenting the results of a VCS project or projects
16	Search Error	Incomplete or missing records, including books with no reference to chapter or section
18	Social Media	Publications detailing scraping social media for CS data, with passive participation from volunteers
133	Specific Projects	Publications discussing specific VCS projects
19	Specific Tools/Technology	Publications detailing specific tools such as smartphones
39	Use Cases	Use cases of VCS-derived data
7	Volunteer Computing	Publications detailing volunteer computing initiatives

Table 3: Stage one literature review results.

Count	Category	Description
5	Ambiguous	Publications with ambiguous findings which may apply only to offline CS
1	Applications	Publications describing applications of VCS projects or activities
1	Contradiction	Publications which contradicted findings of multiple sampled research papers
8	Duplicates	Repeated publications
9	Inaccessible	Publications which could not be accessed
15	Incomplete/Towards	Publications which represent incomplete or immature research
18	Irrelevant	Publications with no relevance to research topics
5	Little Contribution	Publications lacking a significant contribution or message
1	Methodology	Publications drawing on VCS methodologies
1	No VCS	Publications with no discussion of VCS
61	Offline	Publications which discuss only offline CS
5	Opinion	Publications which represent opinion or editorial pieces
12	Other Crowdsourcing	Publications concerned with other forms of crowdsourcing
8	PhD Thesis	Publications which represent draft or accepted PhD theses
11	Presentation	Publications representing presentations rather than published articles
9	Technical Report	Publications which take the form of a technical report
12	Specific Projects	Publications discussing specific VCS projects
1	Volunteer Computing	Publications detailing volunteer computing initiatives
5	Workshop	Workshop papers

Table 4: Stage two literature review results.

Chapter 4

Characterising Virtual Citizen Science

Before commencing with an examination of the impact of sociality on productivity within Virtual Citizen Science (VCS), it is first necessary to better understand what VCS is. How are VCS projects designed and implemented? What patterns of behaviour are typical within VCS? What factors motivate participation? What form does sociality take? It would be impossible to identify the impact of sociality in VCS without an understanding of these questions. Thus, this chapter addresses research question 1: *how is Virtual Citizen Science understood and realised today?* As illustrated in figure 5, this step used a systematic review of VCS literature to understand the state of VCS research, the methodology of which is described in section 3.5.

This chapter is structured as follows. In section 4.3, I outline the sixty one publications identified and analysed through the literature review process, followed by a summary of the projects extracted from these publications in Section 4.4. The main findings and outcomes are then synthesised and presented across four main themes. Section 4.5 outlines significant design challenges associated with VCS within the literature and solutions and features that have been proposed to address these issues. A central element of VCS is the use of *volunteer* effort, rather than paid participants, and the factors that motivate participation and format that this activity takes are core to VCS as a methodology. Sections 4.6 and 4.7 therefore outline the motivational factors and engagement patterns associated with VCS projects. Finally, it is necessary to understand sociality and social features within VCS projects, which is addressed in section 4.8. A

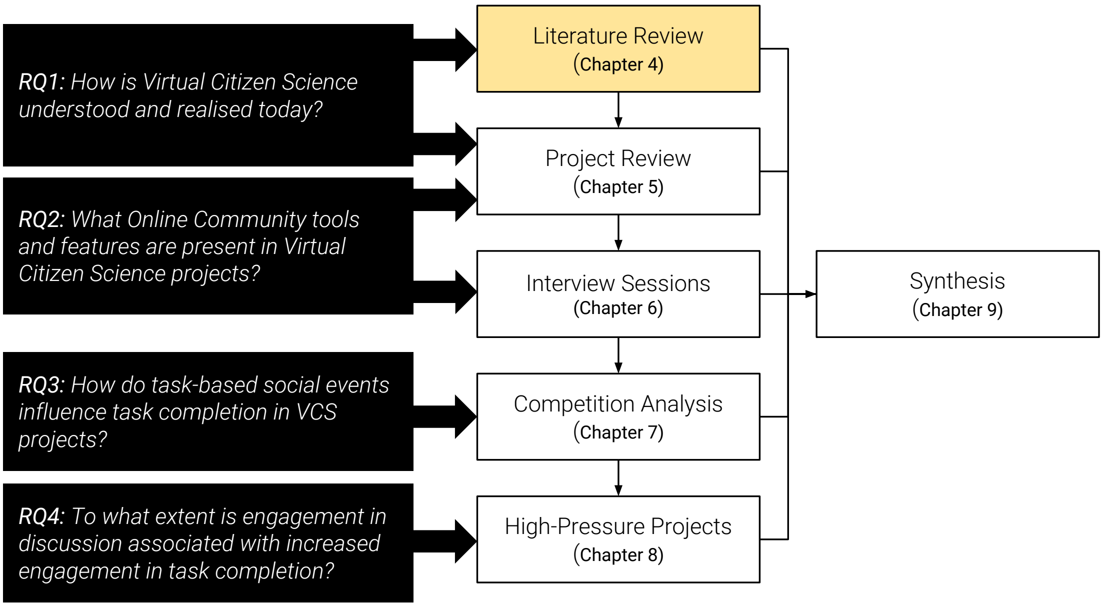


Figure 5: The first study used a systematic review to understand current research on Virtual Citizen Science.

full list of all the publications identified during the literature review process and the categories to which they were assigned can be seen in appendix A and B.

4.1 Further Clarification of Scope

As demonstrated by chapter 2, there are implicit and explicit contradictions within the literature with regard to the definition of Virtual Citizen Science. The definition described by Wiggins and Crowston (2011) of an online only, active process contradicts the citizen sensing and volunteer computing categories defined by Haklay (2013) and also the data collection category described by Tinati et al. (2015b). Similarly, the social machines paradigm and its applicability to VCS strongly contradicts the view of VCS as a passive or administrative process on the part of volunteers – such as is the case in citizen sensing or volunteer computing. In turn, publications identified during the literature process described an 'unaware' Citizen Science process where statements or data published by unwitting members of the general public on social media channels are harvested and used for scientific purposes – a process which strongly contradicts aspects of each of these paradigms, particularly in terms of conscious volunteer-based and volunteer-led activities.

It would be impossible to consider each of these competing definitions and a single, working definition of VCS is essential to the accuracy of the work conducted within this research, including within the literature review described in this chapter. On this basis and for the sake of this thesis, the working definition of VCS will be that described by Wiggins and Crowston – an **active, crowd-sourced** process almost entirely conducted **within a virtual space** such as a website, **designed explicitly for such a purpose**, in which all **participants knowingly partake** on a **volunteer basis**. This is also the definition that best provides for and limits the scope of sociality within VCS initiatives. In particular, offline and social media based channels raise the possibility of unobservable interactions while passive processes such as volunteer computing and citizen sensing strongly limit opportunities for interaction between sociality and task completion.

4.2 Theme Selection

The aim of this literature review is to identify how Virtual Citizen Science is understood and realised today. To that end, the literature sources have been selected and categorised into common themes to understand separate aspects of the characteristics and implementation of VCS initiatives. As previously identified, irrelevant sources or findings which were highly contradictory or did not deal with VCS have been removed from this review. Sources removed in this way can be seen in appendix A and appendix B.

The first of these themes describes the types of data available to volunteers and the types of activity that they must complete within VCS tasks, as well as the domains associated with these projects. This theme is particularly important to conceptualise and describe commonalities within VCS projects and to expand on the characteristics described in chapter 2 with regard to the limits of VCS and how it can be distinguished from other citizen science activities.

Following this, it is necessary to consider the design challenges that are associated with VCS initiatives. This second theme summarises the specific issues that are the focus of wider research in terms of the design and implementation of VCS crowdsourcing platforms. Each of these design challenges has been decisive in shaping the format and features associated with VCS, while also serving to contextualise the broader literature considered and issues raised throughout this thesis.

If the aim of this thesis is to consider a potential relationship between social activity and productivity in VCS, then it is crucial to identify the factors that cause volunteers to engage in VCS projects. The third theme identifies the most commonly cited motivational factors within the literature from across projects, arranged into categories *based on the terminology* within the wider literature. This is followed by a summary of engagement patterns and behaviour within VCS, to understand how these motivations translate into volunteer activity and to serve as a baseline with which to compare findings surrounding increased productivity.

Finally, this review outlines the role and nature of sociality and opportunities for interaction within VCS, as outlined by the literature. In particular, the theme considers the forms and types of interaction available to volunteers within VCS and the platforms that facilitate this interaction, the motivations associated with this interaction (to the extent that they differ from the motivation to participate at all in VCS) and the links suggested between task completion and discussion activity. Ultimately, this theme serves to identify the related work and research conducted so far within this area, but also as a means with which to validate and compare later findings and to ensure that later analysis does not stray outside the bounds of VCS interaction.

4.3 Literature Results

In total 61 publications adhered to the inclusion criteria set out in chapter 3 and were therefore selected for analysis and review. Table 5 includes a full list of each of the sampled publications, along with the source (either conference or journal paper), methodology (and where appropriate sample size) and main contributions of the paper. A small number of authors are listed multiple times, but these publications represent distinct research topics and findings with relevance and significance to VCS. Most notably, Nov et al. (2011a), Nov et al. (2011b) and Nov et al. (2014) are ambiguous but appear to use the same survey results given the low likelihood that an identical sample size and response rate would occur in three distinct studies of the same project. Since each of these publications makes distinct contributions and comparisons with regard to other projects and the wider field, all three publications were included in the final literature sample.

Table 5: Results of systematic review of Virtual Citizen Science literature.

Reference	Publication Source	Methodology (Sample Size)	Key Findings
Baruch et al. (2016)	Journal	Survey (N=2329, 166, 188), Forum Observation (N=60), Interview (N=6, 1)	Motivation (Altruism, Learning), Barriers (Poor asset quality)
Bauer and Popovic (2017)	Conference	Quantitative Analysis - Task Submissions (N=179,723)	Design - Contributions, Accuracy
Cappa et al. (2016)	Journal	Experiment (N=85)	Motivation (Learning), Sociality
Cox et al. (2015c)	Journal	Quantitative Analysis - Task Submissions, Outputs	Evaluation
Cox et al. (2015a)	Journal	Survey (N=1985), Quantitative Analysis - Task Submissions	Motivation
Crall et al. (2017)	Journal	Web-based Experiment (N=254,101)	Design - Recruitment, Retention
Crowston et al. (2018)	Conference	Case Study	Design - Facilitating
Curtis (2015a)	Journal	Forum Observation, Task Observation, Survey (N=37), Interview (N=10)	Motivation (Altruism, Intrinsic)
Darch (2017)	Journal	Case Study	Design - Recruitment, Encouraging Contributions
Desell et al. (2015)	Journal	Web-based Experiment (N=456)	Design - Encouraging Contributions, Accuracy, Recruitment
Diner et al. (2018)	Journal	Experiment (N=140)	Design - Encouraging Contributions
Eveleigh et al. (2014)	Conference	Survey (N=299), Quantitative Analysis - Task Submissions	Motivations, Engagement
Fernandez-Marquez et al. (2016)	Journal	Quantitative Analysis - Task Submissions	Motivations, Engagement
Greenhill et al. (2014)	Journal	Virtual Ethnography, Interview (N=16), Content Analysis	Motivations, Engagement, Design - Contributions
Gugliucci et al. (2014)	Conference	Survey (N=334)	Motivations (Altruism, Intrinsic)
Iacovides et al. (2013)	Conference	Interview (N=8)	Motivations (Intrinsic), Design - Gamification
Jackson et al. (2014)	Conference	Interview (N=3)	Motivation, Design - Retention, Sociality
Jackson and DeVries (2015)	Conference	Quantitative Analysis - Task Submissions	Design - Recruitment, Retention

(Continued)

Reference	Publication Source	Methodology (Sample Size)	Key Findings
Jackson et al. (2016a)	Conference	Web-based Experiment (N=3,553)	Design - Recruitment, Contributions
Jackson et al. (2016c)	Conference	Web-based Experiment (N=6,354)	Design - Contributions
Jackson et al. (2016b)	Conference	Virtual Ethnography	Design - Recruitment, Retention
Jackson et al. (2017)	Conference	Quantitative Analysis - Task Submissions, Interview (N=4)	Design - Facilitating
Jackson et al. (2018)	Conference	Trace Ethnography	Design - Contributions
Jay et al. (2016)	Conference	Web-based Experiment (N=1511)	Design - Contributions
Jennett et al. (2013)	Conference	Interview, Participant Observation	Sociality
Jennett et al. (2014)	Journal	Interview (N=8), Focus Group	Motivation
Kaufman et al. (2016)	Conference	Experiment (N=141)	Motivation, Design - Recruitment, Contributions
Kosmala et al. (2016)	Journal	Literature Review	Design - Accuracy
Laut et al. (2016)	Journal	Experiment (N=120)	Design - Contribution
Lee et al. (2017)	Conference	Experiment (N=36,513)	Design - Recruitment
Lee et al. (2018)	Journal	Web-based Experiment (N=3616)	Design - Recruitment, Motivation
Luczak-Roesch et al. (2014)	Conference	Content Analysis, Quantitative Analysis - Sociality	Sociality
Mao et al. (2013a)	Conference	Web-based Experiment (N¿200)	Design - Contribution
Mao et al. (2013b)	Conference	Virtual Ethnography	Engagement, Design - Contribution, Retention
Morais et al. (2013)	Journal	Quantitative Analysis - Task Submissions (N=146,669)	Engagement Design - Recruitment, Retention
Mugar et al. (2014)	Conference	Virtual Ethnography	Sociality
Mugar et al. (2015)	Conference	Virtual Ethnography, Interview (N=21)	Sociality

(Continued)

Reference	Publication Source	Methodology (Sample Size)	Key Findings
Nov et al. (2011a)	Conference	Survey (N=139)	Motivation (Collective)
Nov et al. (2011b)	Conference	Survey (N=139)	Motivation, Engagement
Nov et al. (2014)	Journal	Survey (N=139)	Motivation
Ponti et al. (2018)	Journal	Virtual Ethnography	Motivation, Sociality
Prestopnik et al. (2014)	Conference	Quantitative Analysis - Task Submissions (N=900)	Engagement, Design - Accuracy
Prestopnik et al. (2017)	Journal	Quantitative Analysis - Task Submissions (N=4174)	Design - Recruitment, Retention, Accuracy
Raddick et al. (2013)	Journal	Survey (N=10,991)	Motivation (Intrinsic, Altruism)
Reed et al. (2013)	Conference	Survey (N=199)	Motivation
Rotman et al. (2012)	Conference	Survey (N=142)	Motivation (Collectivism, Altruism)
Rotman et al. (2014)	Conference	Interview (N=44)	Motivation
Sauermann and Franzoni (2015)	Journal	Quantitative Analysis - Task Submissions	Engagement
Segal et al. (2015)	Conference	Survey (N=257), Quantitative Analysis - Task Submissions, Interview (N=1,679)	Design, Retention
Spiers et al. (2018)	Conference	Quantitative Analysis	Engagement
Sprinks et al. (2014)	Conference	Experiment (N=30)	Design - Accuracy
Sprinks et al. (2017)	Journal	Experiment, Survey (N=30)	Design - Accuracy
Tinati et al. (2014)	Conference	Quantitative Analysis - Task and Sociality (N=2928)	Engagement, Sociality
Tinati et al. (2015a)	Conference	Quantitative Analysis - Task and Sociality (N=98,224)	Engagement, Sociality
Tinati et al. (2015b)	Conference	Interview (N=2)	Design
Tinati et al. (2017b)	Conference	Survey (N=1365), Virtual Ethnography	Sociality, Motivation

(Continued)

Reference	Publication Source	Methodology (Sample Size)	Key Findings
Tinati et al. (2017a)	Journal	Survey (N=1505)	Motivation (Altruism)
Heaton and Torres (2015)	Journal	Virtual Ethnography, Interview (N=8)	Design - Engagement
Wald et al. (2015)	Journal	Platform Survey	Design - Recruitment, Retention
Woodcock et al. (2017)	Journal	Interview (N=23, 19), Virtual Ethnography	Motivation
Zhou et al. (2017)	Conference	Experiment (N=506), Survey (N=100)	Design - Recruitment

4.4 Tasks and Assets

Table 6: VCS projects identified from the literature review sample.

Name	Status	URL	Domain	Asset	Task
Ancient Lives	Active	`https://www.ancientlives.org`	Archaeology, Humanities	Image	Linking, Transcription
Andromeda Project	Retired	`https://www.andromedaproject.org`	Astrophysics	Image	Mapping
Artigo	Active	`https://www.artigo.org`	Art History, Humanities	Image	Cataloguing
Asteroid Zoo	Retired	`https://www.asteroidzoo.org/`	Astrophysics	Image	Mapping
Bat Detective	Active	`https://www.batdetective.org/`	Biology, Ecology	Sound/Graph	Categorising, Contextualising, Mapping
Brooklyn Atlantis	Ambiguous	`www.brooklynatlantis.poly.edu`	Hydrology	Image	Categorising
CamClickr	Retired	`http://watch.birds.cornell.edu/CamClickr/`	Ecology, Ornithology	Image	Categorising
Cell Slider (AKA Click-To-Cure)	Retired	`https://www.cellslider.net/`	Biology, Oncology	Image	Categorising
Chicago Wildlife Watch	Active (Relaunch)	`https://www.zooniverse.org/projects/zooniverse/chicago-wildlife-watch`	Ecology	Image	Categorising
Citizen Sky	Retired	N/A (No longer available)	Astrophysics	Unknown	Unknown
Citizen Sort	Active	`https://citizensort.org/`	Biology, Ecology, Computer Science	Image	Linking, Categorising
Condor Watch	Active	`https://www.condorwatch.org/`	Biology, Ecology	Image	Cataloguing
CosmoQuest	Active	`https://cosmoquest.org/`	Astrophysics	Image	Mapping

(Continued)

Name	Status	URL	Domain	Asset	Task
Cropland Capture	Active	https://www.geo-wiki.org	Various	Image	Categorising
Crowcrafting	Active	https://crowdcrafting.org/	Various	Various	Various
Cyclone Center	Active	https://www.cyclonecenter.org/	Meteorology	Image	Categorising, Linking
Disk Detective	Active	https://www.diskdetective.org/	Astrophysics	Image	Categorising
EteRNA	Active	https://eternagame.org/web/	Biology	Game	Linking, Puzzle Solving
EyeWire	Active	https://eyewire.org/explore	Biology	Game	Mapping
Floating Forests	Active (Relaunch)	https://www.zooniverse.org/projects/zooniverse/floating-forests	Biology	Image	Mapping, Categorising
FoldIt	Active	https://fold.it	Biology	Game	Puzzle Solving
Forgotten Island	Active	https://citizensort.org/web.php/forgottenisland	Biology, Ecology	Image	Linking
Galaxy Zoo	Active	https://www.zooniverse.org/projects/zookeeper/galaxy-zoo/	Astrophysics	Image	Categorising
Galaxy Zoo Bar Lengths	Retired	https://www.zooniverse.org/projects/vrooje/galaxy-zoo-bar-lengths	Astrophysics	Image	Categorising
Galaxy Zoo Hubble	Retired	http://hubble.galaxyzoo.org	Astrophysics	Image	Categorising
Galaxy Zoo Mergers	Retired	http://mergers.galaxyzoo.org/	Astrophysics	Image	Categorising

(Continued)

Name	Status	URL	Domain	Asset	Task
Galaxy Zoo Quench	Retired (Unsuccessful)	`https://quench.galaxyzoo.org`	Astrophysics	Image	Categorising
Galaxy Zoo Supernova	Retired	N/A	Astrophysics	Image	Categorising
Gravity Spy	Active	`https://www.zooniverse.org/projects/zooniverse/gravity-spy`	Astrophysics	Graph	Categorising
Higgs Hunters	Active	`https://www.higgshunters.org/`	Particle Physics	Graph	Mapping
Ice Hunters	Retired	N/A	Astrophysics	Image	Categorising
iSpot Nature	Active	`https://www.ispotnature.org/`	Ecology	Image	Creating Content, Categorising
Microplants	Active	`http://microplants.fieldmuseum.org/`	Biology	Image	Mapping, Cataloguing
Microscopy Masters	Retired	`https://www.zooniverse.org/projects/jbrugg/microscopy-masters`	Biology	Image	Mapping
Moon Zoo	Retired	`https://www.moonzoo.org/`	Astrophysics	Image	Mapping
NEEMO	Retired	`https://neemo.zooniverse.org/`	Astrophysics	Image	Categorising
Notes from Nature	Active	`https://www.notesfromnature.org/`	Various (Predominantly Biology)	Image, Text	Transcription
Old Weather	Active (Relaunch)	`https://www.oldweather.org/`	Meteorology	Text	Transcription
Operation War Diary	Active	`https://www.operationwardiary.org/`	History	Text	Transcription
Penguin Watch	Active	`https://www.penguinwatch.org`	Ecology	Image	Categorising, Cataloguing

(Continued)

Name	Status	URL	Domain	Asset	Task
Phylo	Active	https://phylo.cs.mcgill.ca/	Biology	Game	Puzzle Solving, Linking
Planet Four	Active	https://www.planetfour.org/	Astrophysics	Image	Mapping
Planet Four Craters	Active	https://craters.planetfour.org/	Astrophysics	Image	Categorising, Mapping
Planet Hunters	Active (Relaunch)	https://www.planethunters.org/	Astrophysics	Graph	Mapping
Plankton Portal	Active	https://www.planktonportal.org/	Biology	Image	Mapping
Quantum Moves	On Hold	https://www.scienceathome.org/games/quantum-moves/	Physics	Game	Puzzle Solving
Radio Galaxy Zoo	Active (Relaunch)	https://radio.galaxyzoo.org/	Astrophysics	Image	Mapping
Scistarter	Active	https://scistarter.com/	Various	Various	Various
Seafloor Explorer	Retired	https://www.seafloorexplorer.org/	Marine Biology	Image	Categorising, Mapping
Season Spotter	Active	http://seasonspotter.org/	Biology	Image	Categorising, Mapping
Snapshot Serengeti	Active (Relaunch)	https://www.zooniverse.org/projects/zooniverse/snapshot-serengeti	Ecology	Image	Categorising
Solar Stormwatch	Retired*	N/A	Astrophysics	Image, Video	Classifying, Mapping, Categorising
SpaceWarps	Active (Relaunch)	https://spacewarps.org	Astrophysics	Image	Categorising, Mapping
Stardust@Home	Active	http://stardustathome.ssl.berkeley.edu/	Astrophysics	Image, Video	Categorising, Mapping

(Continued)

Name	Status	URL	Domain	Asset	Task
Sunspotter	Active	https://www.sunspotter.org/	Astrophysics	Image	Categorising
Supernova Hunters	Active	https://www.zooniverse.org/projects/dwright04/supernova-hunters	Astrophysics	Image	Categorising
The Milky Way Project	Active	https://www.zooniverse.org/projects/povich/milky-way-project	Astrophysics	Image	Mapping
Tomnod	Active	https://www.tomnod.com/	Various	Image	Various
Transcribe Bentham	Active	http://www.transcribe-bentham.da.ulcc.ac.uk/td/Transcribe_Bentham	Art History	Text	Transcription
Virtual Atom Smasher	Active	http://test4theory.cern.ch/about/	Particle Physics	Game	Puzzle Solving
Whale FM	Retired	https://whale.fm/	Biology	Sound	Linking
Wildcam Gorangosa	Active	https://www.wildcamgorongosa.org	Ecology	Image	Categorising
Wildlife@Home	Active	https://csgrid.org/csg/wildlife/	Ecology	Video, Image	Categorising, Mapping
Worm Watch Lab	Active (Relaunch)	https://www.wormwatchlab.org/	Biology	Video	Mapping
ZenTag	Active	https://tiltfactor2.dartmouth.edu/zentag	Humanities	Image	Collaborative Tagging
Zooniverse	Active	https://www.zooniverse.org/	Various	Various	Various

Upon completion of the sampling process, each publication was reviewed to identify any VCS projects described or analysed within each source. A total of 66 projects were found to fit the criteria outlined previously by Wiggins and Crowston (2011). These can be seen in Table 6, along with the project status (active or retired), URL (if still available), scientific field as described in the literature sources, asset type and task processes. As there is no common vocabulary for describing VCS tasks, these are described based on the typology outlined by Dunn and Hedges (2013), which is intended to cover digital humanities crowdsourcing, but which adequately describes and defines all of the task processes used within these projects:

1. Cataloguing - Adding metadata to describe and explain assets. A more free process than categorising.
2. Categorising - Assigning assets to one or more categories.
3. Collaborative Tagging - Working alongside another player to offer categorising tags to an asset at the same time, generally in a game-based manner.
4. Contextualising - Adding metadata to place assets within a given context. More involved and complex than cataloguing.
5. Linking - Drawing links between two or more assets of a similar nature or with similar content.
6. Mapping - Identifying structures within images and tracing these onto the asset.
7. Transcription - Digitising text from images or physical pages.

In addition to this, I also add the *puzzle solving* task type, which refers to gamified projects where the asset used is a puzzle or puzzle-like task. These activities do not map well to the other task types described by Dunn and Hedges (2013), although they could potentially be seen as reviewing tasks where the participant reviews solutions so that they can be built on by an algorithm.

The iSpot Nature project is a notable exception within this list, as it features both investigative and virtual features, encouraging participants to take images of wildlife in their local area, before uploading these to the platform for other volunteers to classify. Since it is possible for participants to contribute entirely through a virtual portal and in an active manner, the project was deemed to meet the criteria of a VCS project and was therefore also included. Many of the projects had been retired and relaunched – where this has happened, the information included represents the project as it stood before being relaunched. This was because the publications generally dealt with the project in its previous form, or did not offer sufficient information about the relaunched project to offer entirely up-to-date details or to identify any differences. There was little information about Citizen Sky within the literature and no web presence remained for the project and it was therefore not possible to ascertain the assets or task processes involved within the project – the project is included here simply in the interests of completeness.

Despite the apparent diversity of Citizen Science as a field as suggested by Kullenberg and Kasperowski (2016), the projects mentioned within the sampled literature largely share a small number of specific common features. In terms of scientific domain, just nine domains were represented by the projects mentioned within each publication. The majority of these (25 projects) correspond to astrophysics, predominantly from the Zooniverse platform, which is perhaps unsurprising given that Galaxy Zoo was a pioneering project within VCS and Zooniverse volunteers have shown the tendency to contribute to projects from the same domain (Luczak-Roesch et al., 2014). The second largest group (18 projects) corresponds broadly to the biological sciences – either ecology (11), medicine (6) or marine biology (1). Beyond the two largest domains, there are also projects from the humanities generally and more specifically, history and art history, as well as particle physics and hydrology. On the one hand, these results are clearly skewed towards the chosen methodology, which in drawing on peer-reviewed, published research is partially incompatible with the slow rate of publication associated with Citizen Science projects (Theobald et al., 2015). However, indications from the Zooniverse platform indicate just one further domain not identified using the literature review process – social science, as well as the "language" category as a specific form of humanities-based VCS research.

With regard to assets, six varieties were identified - images, video, text, graphs, sound recordings and game-based activities. Despite this, the vast majority of projects used images, including automated camera-trap images, satellite imagery and images sourced from scientific equipment such as the Sloan Digital Sky Survey. Six projects used a game-based interface, although each project used a different, custom interface from more gamified, puzzle-based activities in the case of EteRNA and Phylo, to more complex and less clearly game-based activities in the case of EyeWire and FoldIt. Four projects used exclusively text-based assets for transcription tasks, which were distinct from images, as although each was delivered to participants in the form of an image, the underlying asset in these cases represented a physical piece of paper, such as a ship log-book, diary or museum exhibit record. Just two projects use audio recordings and these are both earlier Zooniverse projects (Tinati et al., 2015b). Four projects used videos, but in the case of Stardust@Home, the function of the video was predominantly to group together several images to correct issues with the resolution stemming from the microscope used to collect asset images and so these were not videos in the same sense as those used within the other three projects. A further four projects used graphs representing data recorded from astrophysical phenomena, although there is no reason to believe that such assets could not be applied in other contexts.

Each of the projects only corresponded to the contributory category outlined in chapter 2, with participants contributing only to the data analysis process. EteRNA advertises the opportunity for volunteers to contribute to scientific publications, but this was not mentioned within the sampled publications. One notable exception to this was Galaxy Zoo Quench, which intended to offer participants the opportunity to contribute to all stages of the scientific research process, gathering data (by analysing assets), analysing the results of this process and then preparing and publishing these findings, similar to a collaborative project (Crowston et al., 2018). Unfortunately, however, the project did not proceed beyond the second phase – analysing the results of the crowdsourcing process – largely due to the difficulty of the task and the lack of explicit management and leadership among the project participants (Crowston et al., 2018). This suggests that it may be extremely difficult for a VCS project to successfully employ a collaborative or extreme Citizen Science methodology, except in cases where the size of the active community is very small. Moreover, the Galaxy Zoo Quench project was far from an extreme Citizen Science project, as participants were provided with a pre-defined research question and materials, selected to facilitate crowdsourcing processes (Crowston et al., 2018).

4.5 Design Challenges

The majority of the sampled publications dealt with empirical observations and experiments with regard to how best to design and implement VCS pro-

Figure 6: *About* page from the EteRNA project, advertising the opportunity for participants to suggest experiments and contribute to research publications

jects surrounding four design challenges: accuracy and data quality; maximising contributions and productivity; participant recruitment and finally participant retention. Although not all of these challenges directly relate to issues of productivity, many of the solutions proposed have a direct effect on the workload expected of volunteers and a knock-on effect in terms of the importance of maximising volunteer efficiency and effort.

4.5.1 Accuracy and Data Quality

Maximising the accuracy of submissions to VCS projects is an essential and fundamental area of design-based research. In fact, one of the factors motivating the use of human-in-the-loop processes within VCS is the low performance and lack of accuracy associated with algorithms for processes such as protein folding and morphological galaxy classifications (Khatib et al., 2011; Lintott et al., 2008). While VCS carries with it many of the same biases and data quality issues present in professional research, projects must also contend with biases specific to Citizen Science, particularly a "high variability... of [volunteer] demographics, ability, effort, and commitment" (Kosmala et al., 2016). Similarly, projects often deal with unfamiliar concepts or expect participants to engage with new tools in a way with which they are not familiar (Nov et al., 2014).

If classifications are to be used in scientific research, then project science teams need to know that the classifications submitted by players are as accurate as possible. On the whole, however, accuracy and data quality issues are not major factors connected to productivity and sociality within VCS, particularly since solutions to data may be addressed through changes to task design or the application of different algorithmic models to weight and aggregate responses, without the awareness of project participants (Antelio et al., 2012). Still, there are indirect connections between these issues and productivity, as data quality issues may motivate the introduction of additional tasks such as community review or redundancy in turn increase the workload expected of players and heighten issues posed by the skewed activity patterns within VCS. For this reason, these issues are presented here to better situate issues related to productivity within associated design issues.

VCS methodologies are associated with redundancy and aggregation, as multiple participants complete the same task for the same asset before a weighted or simple majority is calculated to increase overall accuracy. Due to the methodology used to conduct this literature review, it is not possible to state with absolute certainty whether any given quality assessment process is used within any given project, as papers offering only details about the methods used in specific projects were rejected on the grounds of low relevancy and few of these publications offered significant details about quality assessment mechanisms. Wiggins et al. (2011) suggest that 23% of sampled projects made use of redundancy and aggregation, although notably this study included offline CS projects and was conducted in 2011, at an early stage of the boom in Citizen Science projects that has been experienced in recent years (Kosmala et al., 2016). Regardless, in their original definition of Virtual Citizen Science as a classification, Wiggins and Crowston (2011) identified redundancy as a defining and common feature of the then emergent area and redundancy is used across Zooniverse projects, although the specific details of the aggregation vary from project to project[1].

For the purposes of this thesis, the specific quality assurance mechanisms used within projects are not necessarily significant, but the use of redundancy carries with it increased overall project workloads, as the number of tasks that must be completed by the community increases. In Snapshot Serengeti, for example, assets marked with the "nothing here" tool are shown to 25 differ-

[1]https://blog.zooniverse.org/2013/06/20/how-the-zooniverse-works-the-domain-model/

ent users Swanson et al. (2015). Furthermore, task design policies such as the individual-focused workflows described by Mugar et al. (2014) may drive participants to engage with social features. Another key finding from the literature is that projects may further increase the workload assigned to players through the use of simulated, tutorial or gold-standard images, previously classified by experts for calibration purposes, but these have been associated with reduced motivation from participants and the potential waste of valuable participant effort that could be better assigned to 'real' tasks (Darch, 2017; Ponciano and Brasileiro, 2015; Tinati et al., 2015b). These gold standards can be shown to participants as part of task workflows, as in Planet Hunters where simulated assets make up 5% of tasks, but can also be applied by experts after data have been collected or for accuracy estimation purposes, without further work from participants (Fischer et al., 2012; Swanson et al., 2015). Evidence from the literature in this regard was insufficient to identify if these practises are widespread and which approach, if any, projects use.

Issues of task design and the presentation of assets have significant impacts on the quality of data generated by VCS projects even within the same project and community. In Wildlife@Home, increasing the length or completeness of video assets has been demonstrated to lead to increased engagement and accuracy by offering increased context to volunteers which renders tasks easier for participants to complete, although accuracy diminishes sharply for video assets in excess of 5 minutes (Desell et al., 2015). Similar findings with regard to engagement in text transcribing tasks have been suggested by Eveleigh et al. (2013) and Tinati et al. (2015b), as participants became immersed in the narratives that could be generated from contextual details, but no impact on overall accuracy or ease-of-completion was described in these projects. Conversely, game-based tasks with integrated narratives do not significantly impact data quality, but do lead to an increase in cheating and low-effort behaviour designed to game the system and achieve maximum progress and rewards with minimum effort (Prestopnik et al., 2014). At the same time, it is unclear to what extent this is a result of the game-based experience itself and to what extent this is simply a result of design decisions such as reward mechanisms.

Perhaps counter-intuitively, Desell et al. (2015) and Sprinks et al. (2017) both suggest that more difficult tasks lead to greater overall accuracy in participant contributions. Difficulty, however, is subjective and both publications make use of participants' opinions regarding task and interface simplicity, rather than more objective measures. There are also clear limits to the extent to which task difficulty is beneficial, as Desell et al. (2015) identify that "medium" difficulty tasks generate the most accurate contributions, yet in the experiment conducted by Sprinks et al. (2017), it was the median difficulty task which generated the least accurate submissions, while the easiest interface achieved peak accuracy. More ambiguously still, Kosmala et al. (2016) point to previous findings from the Snapshot Serengeti project suggesting easier tasks lead to the highest accuracy levels, although overall accuracy for newcomers is lower than for more experienced participants. While somewhat contradictory, these findings are significant for this thesis given the possibility raised by Khatib et al. (2011) and particularly by Mugar et al. (2014) that tutorial processes and task completion alone may be insufficient for participants to learn to complete VCS tasks, with Mugar et al. (2014) noting that social features and discussion are essential for participant learning and feedback. Similar findings are suggested by Prestopnik et al. (2014), who note that participants who contribute for longer periods and in greater quantities within Forgotten Island and Happy Match show no statistically significant increase in classification accuracy.

4.5.2 Maximising Contributions

Maximising contribution levels from volunteers is a significant and central issue within VCS (Kosmala et al., 2016). There are strong pressures stemming from the high inequality and general lack of activity demonstrated by many VCS volunteers (see section 4.7) and the extent to which project teams can address individuals' participation patterns and influence volunteer retention are limited, even for the most active and motivated participants (Sauermann and Franzoni, 2015). Three main approaches have been suggested within the literature - gamification and extrinsic motivational factors such as points and prizes, anchoring and the use of messages designed to reinforce participants' intrinsic motivations and sociality-based approaches making use of competition and goal-setting relative to other members of the community. However, it is unlikely that these are the only effective solutions to the issue of maximising contribution patterns, as even apparently insignificant factors can have a large impact on volunteers' contribution levels. For example, offering volunteers the choice between registering for projects and contributing anonymously increases overall participation by as much as 60%, even increasing participation from those volunteers who would contribute regardless of the need to register (Jay et al., 2016). Similarly, factors that would otherwise be left to chance - for example, the content of the initial assets that a volunteer classifies - can have a significant effect on participants' intentions to continue to contribute and as a result, on their overall contribution levels (Mao et al., 2013b; Tinati et al., 2015b).

In an experiment conducted with a "virtual peer" performing based on different algorithmic models, Laut et al. (2016) found that participants will contribute significantly more classifications if presented with indicators to the performance of fellow participants, but only if these fellow participants perform within a certain range. In other words, extremely high performing participants actually have a negative effect on a participant's desire to contribute, leading to slightly lower contribution levels overall, with mid-performing participants having the greatest impact on classification levels. This does not extend to the completeness of each classification, with participants contributing more tags per image in line with indications of fellow participants – although the number of tags contributed is closer to the performance of peers when these peers contribute a low to average number of tags. These results align with interview findings from the Old Weather project, where participants noted that competing with high performing players was a stressful experience that resulted in reduced interest in contributing (Eveleigh et al., 2013). However, as shown in section 4.6.6 competition has been noted to have no effect on - or to even be demotivating for - participants in both non-gamified projects and Games with a Purpose (Curtis, 2015a; Darch, 2017).

Alternatively, the beneficial effect of the virtual peer may not be a social factor at all, but rather an indication of the effectiveness of goals and aims within VCS. In an experiment conducted by Jackson et al. (2016a), participants who set personal classification goals - for example, to classify 50 images in a single session - made on average more than twice as many classifications as participants who chose not to do so. The use of anchoring messages relative to high performing members of the community was found to encourage the setting of higher goals than messages related to one's own prior performance, but did not have a statistically significant effect on the final number of classifications made. Conversely, an experiment within the Higgs Hunters project found that anchoring messages informing participants if they were the first to see a given image resulted in an increase in the average number of sessions and contributions made per session, although the authors caution that approximately

33% of the non-control group are unlikely to have seen these indicators and it is therefore not possible to entirely accurately deduce the impact of discovery-based anchoring. In an experiment by Kaufman et al. (2016) within the Zentag project, participants were found to contribute more tags per classification and were more likely to replay the game when presented with messages regarding the wider community which nonetheless emphasised the opportunity to be the first to classify an image than in messages emphasising social norms, intrinsic motives, altruism or indeed the chance to be one of many who would examine an image. Again, the impact of any social factors within this experiment were ambiguous, as compared to social norms, which generated the least number of tags, the inclusion of messages emphasising the opportunity to contribute alongside the community generated a moderate increase in tagging levels.

The use of gamification has been suggested as a means of increasing engagement and thereby, contribution levels within projects. Zhou et al. (2017) found that the use of game elements such as leaderboards allowed players to view how other players were participating, making participants feel part of a community and enhancing motivations. On the other hand, such game elements were also associated with a reduced sense of player competence. Iacovides et al. (2013) presents findings that associate game features with extended task contribution periods and increased productivity as a result. In contrast, within the Virtual Atom Smasher game, participant contributions followed much the same distribution of effort as associated with other VCS initiatives, suggesting that gamification alone is insufficient for encouraging contributions (Fernandez-Marquez et al., 2016). This is partially supported by Prestopnik et al. (2014), who found that player retention (and thus, contribution levels) were significantly lower overall in a more narrative-driven, game-based experience than in a simple matching activity with game elements. Games and game elements can also, in some circumstances, have little to no relevance to the scientific activity in question and thus divert effort from more important tasks (Greenhill et al., 2014).

4.5.3 Participant Recruitment

If an online community is to survive, then it is essential that it continuously attracts new members and contributors to overcome user attrition (Kraut et al., 2012). This is particularly the case in VCS projects where churn rates and inequality of effort are extremely high (Nov et al., 2014; Sauermann and Franzoni, 2015). However, while Kraut et al. (2012) suggest that many online communities rely predominantly on word-of-mouth, such an approach is not suitable for VCS, where participants need nudging to discuss projects on social media and are unlikely to discuss their VCS activity in conversations with colleagues and friends (Cox et al., 2015a; Eveleigh et al., 2014; Tinati et al., 2015a). Moreover, in VCS projects, participants recruited during word-of-mouth campaigns tend to contribute fewer classifications and for less time than participants recruited in other methods (Morais et al., 2013). Factors which drive volunteer participation in VCS projects are also still relatively poorly understood, with diverse volunteer and demographic profiles and somewhat contradictory study findings, particularly with regard to the motivations as identified by the volunteers themselves. For example, Galaxy Zoo volunteers as noted are primarily driven by a desire to assist science, but what this actually *means* to the volunteers and what they perceive to align to these motivations are not necessarily the same as what scientists and designers expect (Darch, 2017; Raddick et al., 2009b). It is therefore perhaps unsurprising that within the sampled literature, there has been little consideration for explicitly designing for the recruitment of volunteers beyond understanding and exploiting initial motivations – particularly intrinsic motivations.

Even so, exploiting these motivations is a more complex process than may initially be assumed. In an experiment by Lee et al. (2017) making use of recruitment emails based on four key motivations – learning, contributing to science, community and altruism – there was only one percentage point difference between the most popular message (learning) and the least popular (altruism) and no recruitment message was more than 6% successful. In a similar experiment by Kaufman et al. (2016) in the ZenTag project, the framing of original recruitment messages had no statistically significant influence on the number of participants recruited. Among those participants who do choose to contribute to a project, there is a strong cross-project effect where participants are much more likely to contribute to projects from the same platform within the same or similar domain - for example, moving from one astrophysics project to another (Luczak-Roesch et al., 2014).

4.5.4 Participant Retention

The extent to which project design teams can design for participant retention is a point of some contention within the literature. Sauermann and Franzoni (2015) express the belief that the loss of long-term volunteers is inevitable and results from a variety of projects demonstrate the high user churn rate that is associated with VCS. In Galaxy Zoo, by 2013 just 3.29% of participants had contributed to the project for over a year (Morais et al., 2013) and Ponciano et al. (2014) note that 31% of Galaxy Zoo participants display an extremely low pattern of engagement, contributing for on average just 71 seconds and tending not to return to projects after initial contributions. In FoldIt, most participants do not complete the tutorial process and a similar phenomenon is seen in EyeWire (Curtis, 2015a; Tinati et al., 2015a). Among those participants who leave projects, some may simply forget about the project (Eveleigh et al., 2014). Marketing campaigns through social media and email have been found to reduce the average delay between classification sessions among newcomers to Season Spotter and drive 45% of inactive volunteers to re-engage with projects (Crall et al., 2017). Yet still, even long-term participants to VCS initiatives tend to contribute unpredictably and sporadically, in the case of EyeWire spending longer between tasks and not making contributions than they do actually tracing neurons (Morais et al., 2013; Ponciano et al., 2014; Tinati et al., 2015a).

Retention is nonetheless crucial for the success of VCS projects (Wald et al., 2015). As Kraut et al. (2012) state, long-term members of any online community "work harder, say more, do more" and are willing to assist the community with future problems that it may face. This is no less true in VCS, where overall it is long-term participants who do the majority of the work[2] and who may take on additional responsibilities without request from project science teams (Darch, 2017).

At the same time, task design factors play an important role in the user experience and the subsequent decision by volunteers to continue to contribute to projects. In a survey of 561 Tomnod participants, 100 respondents stated they decided to leave the project based on the poor quality of images, but the second largest group – 80 participants – left the project due to technical issues

[2]Studies have demonstrated this phenomenon within the Zooniverse platform (Luczak-Roesch et al., 2014; Ponciano et al., 2014), FoldIt (Curtis, 2015a) and EyeWire (Tinati et al., 2015a) and a similar observation has been made in other related forms of crowdsourcing (see for example: Aristeidou et al. (2017))

stemming from difficulties with the project website (Baruch et al., 2016). The opportunity to participate as part of a team or group has been suggested by FoldIt players as a key factor in their continued participation (Curtis, 2015a; Iacovides et al., 2013). Game-based elements and elements of little scientific value which are nonetheless fun and which bring the community together – such as image captioning campaigns and competitive games such as Pong – also lead to increased commitment and loyalty on the part of participants (Greenhill et al., 2014). Similarly, participants who engage in chat and social features tend to remain with projects longer than those who do not, particularly where this engagement occurs early in a volunteers' participation in a given project (Jackson et al., 2014, 2016b; Tinati et al., 2015a).

4.6 Volunteer Motivations

Sixteen publications from the sampled literature explicitly dealt with the motivations that drive volunteers to contribute to Citizen Science projects. Understanding these motivations is of particular value for the purpose of this thesis, both as an opportunity to understand what drives engagement with social features and also to understand whether sociality drives task activity. On the other hand, each of these publications shares specific methodological characteristics that renders a full argument drawn solely from literature findings insufficient and the motivational factors presented within this section are unlikely to be truly and accurately indicative of the wider citizen scientist population. Each of the sixteen papers relies predominantly on surveys and to a lesser extent interviews with participants and the response rates for these surveys tend to be extremely low. Moreover, engagement with VCS projects is sporadic and highly skewed towards a small proportion of highly active users, with the majority of users contributing for single brief sessions (see 4.7 for further details). This property, when combined with the short-term sampling methods common to these publications raises the possibility that the respondents to such surveys are disproportionately long-term and highly active users. Indeed, this limitation was clearly stated within many of the motivation studies – as Heaton and Torres (2015) state "Given the small number of interviews...we have no pretensions of representativity."

4.6.1 Altruism

Altruistic motives and a related desire to contribute to research were one of the most commonly identified factors across the studies, with high response rates from participants. Over 60% of volunteers in FoldIt described an altruistic desire to contribute to research (Curtis, 2015b). Participants in some studies explicitly made reference to the opportunity to improve the future for society as a whole - in a survey of 1915 Zooniverse users, the opportunity to contribute to worthy causes and benefit society were the highest scoring factors on a 7-point likert scale, with mean scores of 6.05 and 6.38 respectively. This is no less true of specific projects from within the platform. Among Galaxy Zoo users, 22% of survey respondents stated a similar desire (Raddick et al., 2009b), while participants in a follow up survey scored contributing to science at 5.87 and helping with the project at 5.75 out of 7.

This altruism appears to extend beyond simply helping other people. In a survey of 2329 Tomnod participants, among the 17% of respondents who described their motivations for participation, the largest group described a desire to help others and the environment, while in a follow up study, 33% of all demographic groups identified the opportunity to help others as the most significant factor influencing their desire to participate in Tomnod campaigns (Baruch et al., 2016). It is however, unclear from the responses who it is that participants feel they are helping or how they do so. This partially aligns with the findings of Reed et al. (2013) that 'altruistic' motives described by participants who enjoy volunteering are at least partially for personal benefit. In an exploratory factor analysis of motivations for participation in the Zooniverse platform, volunteering and helping science to make oneself feel good or important were more strongly aligned than volunteering for the purpose of making the world a "better place".

Nevertheless, there are contradictions within the literature findings. In an experiment using different motivational messages to encourage participants to contribute to a VCS project, Lee et al. (2018) found that participants receiving an altruism-based message had a significantly lower click-through rate than those receiving a message concerning contributing to science, joining a community or learning about science. The altruism group did, however, have the highest likelihood to contribute upon clicking through to the platform and contributed the second highest number of classifications overall. One possible explanation for this finding is that altruism and related motives are of less significance upon joining projects and instead drive long-term participation. In a study conducted by Rotman et al. (2012) volunteers and scientists from collaborative Citizen Science projects were asked to score 4 motivation groups using a 5-point likert scale: egoism (personal benefit); collectivism (benefit to identified group); altruism (benefit to non-identified group) and principlism (benefit to everyone). Volunteers scored each motive similarly, with average scores between 4.08 and 4.46, with principlism receiving the lowest score. Nevertheless, in follow-up interviews and in other studies, altruism was found to develop further in those participants who make the shift from short-term to long-term patterns of contribution, suggesting that it is altruism that drives sustained participation (Crowston and Fagnot, 2008; Gugliucci et al., 2014; Rotman et al., 2012). These long-term participants, while a small minority, are nonetheless of the greatest importance to the ultimate success of VCS projects. In Galaxy Zoo and the Milky Way Project, these volunteers represented just 13% and 16% of the community, but were responsible for the majority of the working time contributed – 40% and 46% respectively (Ponciano and Brasileiro, 2015) (see 4.7 for further details).

Even so, anecdotal findings from Galaxy Zoo suggest that this altruism is somewhat conditional, with participants driven to leave projects if they perceive work to be unlikely to immediately contribute to science and therefore to be a waste of their time, regardless of the indirect value that such work may have for scientific research (Darch, 2017). Participants require feedback on the value and significance of their contributions to reinforce their altruistic motives and desire to contribute, if they are to continue to engage with projects in the long-term (Eveleigh et al., 2014; Darch, 2017). This feedback is a key affordance and opportunity offered by the discussion forums and blog features offered within VCS projects (Jackson et al., 2014; Darch, 2017), but must be offered at both the community and individual level. Woodcock et al. (2017) quotes a participant who was concerned that given the sheer quantity of classifications made within the project "does my small/meagre contribution really make much of a difference?"

4.6.2 Intrinsic Motives

Intrinsic motives were the most frequently cited factors within the sampled literature, appearing often in describing the underlying rationale behind design choices as well as in responses from volunteers. At the same time, the specific definition of 'intrinsic' factors varied between papers, with some publications separating out fun, an appreciation of the images used within projects or a drive for self-improvement from factors such as an interest in science in general, particular scientific fields, or the underlying project itself. For simplicity and consistency sake, intrinsic motivation as understood within this section and within this thesis in general refers to the definition popularised by Ryan and Deci (2000), in keeping with the majority of the sampled publications.

Fifteen of the publications described some form of intrinsically motivating aspect to VCS projects or VCS participation. Two of these described interest in Citizen Science broadly, five in science and research, five in the specific scientific field represented by a project. In terms of the projects themselves, four described general curiosity about a project or VCS task, while three described an appreciation for or curiosity about the assets and images used in a project. Self-improvement and challenge were also common intrinsic factors, described by five and four publications respectively. Although intrinsic motives appear to be significant for all forms of Citizen Science, these are particularly strong within active, web-based 'data analysis' projects (Nov et al., 2011b).

In contrast to altruistic motives, which grow in significance as participants contribute, intrinsic motives appear to be significant throughout a participant's lifecycle. Jennett et al. (2016) describe a feedback loop where participants' initial decisions to contribute are predominantly driven by their intrinsic interests and curiosity about a project, with these interests subsequently growing over time as participants gain the opportunity to learn about and experience a project. Scientific interests in particular were found to be the most significant factor in attracting volunteers to the CosmoQuest platform (Gugliucci et al., 2014). 24% of respondents described an interest in astronomy, while a similar 24% stated an interest in participating in scientific research, with both interests receiving the greatest proportion of 7-point scores, from over 50% of respondents. Notably, however, discussion features which best influence learning – see sections 4.6.8 and 4.8.2.

Intrinsic motives appear to be significant across projects. Among Galaxy

Zoo volunteers, 46% of respondents described an interest in astronomy, while during a follow-up survey, interest in science was the most highly rated factor, with a mean score of 6.2 out of 7, with an interest in astronomy second with 6.09 out of 7 (Raddick et al., 2013). Similarly, Nov et al. (2011a) analysed three distinct VCS projects and noted that participants rated intrinsic motives as second only to collective motives in terms of their influence on participation across all three initiatives (Nov et al., 2011a). Even in highly gamified contexts, which tend to rely on extrinsic motivations and rewards, intrinsic motivations drive player participation (Curtis, 2015a)- 17% of EyeWire respondents described an intrinsic interest in the underlying science or project, while 20% described the project as fun even without the game or puzzle elements (Tinati et al., 2017a). On the other hand, these initial drives to participate may be relatively weak and as simple as perceiving a project to be an "intriguing" or "nice" idea (Jennett et al., 2014).

Beyond motivating participation, intrinsic motives have also been suggested to play a key role in influencing the quantity of contributions and time spent within projects by participants. In Galaxy Zoo, participants demonstrated a tendency to contribute more classifications and spend longer contributing when the initial images they are given to classify are interesting or rare (Mao et al., 2013b). Nonetheless, this contradicts the findings of Nov et al. (2014), who found that while the strength of intrinsic motivations is positively correlated with contribution quantity, there is a weak negative correlation between this strength and the quality of resulting contributions.

4.6.3 Extrinsic Motives and Rewards

On the whole, discussion of extrinsic motives or rewards within the sampled literature was rare, likely due to the uncommon nature of such rewards within VCS projects (Woodcock et al., 2017). Findings also appear relatively contradictory, even when discussing the same project or context. In Tomnod, 10% of male and 12% of female participants under 50 wished to see the addition of awards for high achieving Tomnod users, although only 1% of females over 50 and 3% of males over 50 agreed (Baruch et al., 2016). Experiences from Galaxy Zoo suggest that participants did not engage with extrinsic motivational factors such as leaderboards and as a result, these were eventually removed from the project (Darch, 2017), yet Ponti et al. (2018) suggest that participants continue to request such features and it is the Galaxy Zoo team, not the community, who actively resist such requests. Even in gamified contexts however, which heavily draw on rewards such as points, participants express little interest in such mechanisms - either describing an explicit disdain in the case of FoldIt (Curtis, 2015a; Heaton and Torres, 2015) or ranking such factors as significantly lower than altruistic, intrinsic and social motives in the case of EyeWire (Tinati et al., 2017a). There also appears to be an implicit assumption in the literature that extrinsic factors are not motivating in VCS, as Nov et al. (2011a) assert that the low ranking assigned by participants to the reputation motive is directly linked to the perception of reputation-based features as reward mechanisms, although explicit evidence or justification for this view is not offered.

4.6.4 Gamification and Game Elements

Similar to extrinsic and reward factors, views on gamification as a motivator for VCS participation within the literature are somewhat contradictory. Curtis (2015a) offers statements and views from highly active FoldIt participants drawn from participant observations and interview that suggest that the game elements are of little interest to them and neither a positive or negative influence on their motivations within the project. Yet a study of FoldIt participants suggest that the drive for high scores and the resulting behaviours that this encourages are demotivating for the majority of players and at odds with the altruistic and intrinsic value of the project (Ponti et al., 2018). Further contrasting are the findings of Iacovides et al. (2013), that while game elements do not provide motivation to engage with FoldIt, they do drive participants to contribute for longer periods as they attempt to improve on their performance. Similarly, while Ponti et al. (2018) suggest that Galaxy Zoo participants perceive the addition of game elements to the project as trivialising the serious research that is being conducted, Greenhill et al. (2014) note that participants have previously engaged heavily with Zooniverse-linked games such as "Voorwerp Pong" which have no scientific value and indeed, have organised and created their own gamified initiatives. It is notable that none of the publications which deal with gamification as a motive offer evidence for their assertions such as the design experiments described previously and so it is not possible to ascertain how participants' opinions and statements correspond to behaviours and classification intentions. At the same time, Curtis (2015a), Jennett et al. (2013) and Tinati et al. (2017b) all offer the view that sociality plays a key role in the perception and enjoyment of game elements, further obfuscating the specific impacts of adding game elements to VCS tasks.

4.6.5 Collective and Community Motives

Collectivism, community-driven norms and the opportunity to be part of a community were relatively common motives, although the effectiveness of these motives differed between sources. In a survey of motivations in *Stardust@Home*, volunteers ranked collective motives as the highest factor influencing their participation, with a mean score of 6.45 out of 7 (Nov et al., 2011a). In a later follow-up, Nov et al. (2014) identified collective motives - advancing common goals as part of the community - as the most significant factor positively influencing both the quantity and the quality of contributions. A participant observation of players within FoldIt identified the success of teams, the wider community and the project as a whole as important reasons for participation (Curtis, 2015a). During a post-observation interview phase, over a third of FoldIt players described interaction with fellow players as being a factor motivating their continued participation, a finding supported by the importance of relationships built by members of the community (Curtis, 2015a). Beyond relationships with fellow volunteers, participants interviewed by Rotman et al. (2014) desired to build lasting trust with scientists, to mentor others and to learn and work towards common goals (Rotman et al., 2014). Indeed, the presence of project scientists and existence of a wider community of interested participants was identified by Jennett et al. (2013) as drivers for the initial decision to engage in VCS projects. Collective motives and the community were key factors in a cluster produced through an exploratory factor analysis of Galaxy Zoo participants (Reed et al., 2013).

Nevertheless, community-based messages were found to result in the second lowest click-through rate in the experiment conducted by Lee et al. (2018) and also resulted in significantly lower contribution and session counts than altruistic, learning-based and science-related messages. This aligns with findings from EyeWire, where just 3% of respondents identified the community as a motivating factor for their participation in EyeWire, 74% of whom described the community as of secondary rather than primary importance (Tinati et al., 2017a). Ultimately this is likely a result of the overall lack of engagement with community features, with such participation being characterised as meta-participation, largely common only in those participants who choose to go beyond mere task participation (Crowston and Fagnot, 2008; Luczak-Roesch et al., 2014; Tinati et al., 2015a). Furthermore, while collective motives were associated with increased contributions, Nov et al. (2011b) noted that *identification* with the Stardust@Home community was weakly negatively correlated with contribution quantity. It appears, then, that at least in terms of participants' perceptions, there is a difference between being part of a project or campaign and being part of a community.

4.6.6 Competition

Competition was cited as a motive in a small number of projects, including Galaxy Zoo, despite the removal of a leaderboard and competitive features in this project at a very early stage Darch (2017); Raddick et al. (2013). Even so, somewhat unlike other motives, competition serves as a double-edged sword. Those participants who report being motivated by competition with fellow participants also report becoming demotivated as they are outperformed by others, suggesting that competition is only motivating when the performance of other participants is deemed to be attainable (Eveleigh et al., 2013). Experiments using virtual peers have demonstrated that participants will make increased numbers of contributions in response to higher performing peers, although this effect diminishes as the virtual peer's performance increases (Diner et al., 2018; Laut et al., 2016). The findings of these studies do not definitively state, however, whether this effect is entirely competitive or a result of a collectivist, norm-oriented desire to do one's part for the community.

Interestingly, competition does not appear to be a motivation in Games with a Purpose. Just 12 EyeWire participants identified competition as a factor motivating their participation within the project, of whom just 3 stated competition to be a primary motivation (Tinati et al., 2017a). Similarly, in the case of FoldIt, Curtis (2015a) notes that participants were not motivated at all by the competitive nature of the game, although there was a certain element of personal competition, as participants strived to do the best for the project and for their teams (Iacovides et al., 2013). This is reflected in the findings of Ponti et al. (2018), who described statements from FoldIt participants that the competitive nature of the game was at odds with the serious research being done and thus, of little importance to players. These findings conflict with the statements cited by Heaton and Torres (2015), however, that participants are spurred on to participate by the competitive nature of the game and that this competition does not spur group, but rather individual performance. Given that a low response rate impacted both Curtis (2015a) and Heaton and Torres (2015) and the small size of the active FoldIt community, it seems likely that these contradictions simply reflect the opinions and experiences of different parts of the player base.

4.6.7 Sociality and Interaction

Overall, survey responses from participants tended to rank sociality and discussion among the lowest motivational factors. A study of Galaxy Zoo determined community interaction to be the lowest rated of nine surveyed themes, with a score of 2.62 from e-mail respondents and 2.84 from forum respondents (Raddick et al., 2013). In Stardust@Home, too, social interaction received the lowest rating, with a mean score of 3.30 out of 7 (Nov et al., 2011a). This aligns with the findings of Cox et al. (2015a), who find that social interaction does not play a role in influencing participation for the most active and committed players, although the survey conducted by the authors considers sociality only in terms of relationships with friends and family members in the 'real' world. It is suggested that social participation diverts from the task assigned to players, who prefer to spend their effort contributing to the task portion of projects. However, Reed et al. (2013) conversely suggest the reason for this low ranking is partially due to the low proportion of players who actively engage with discussion features. Once again, this aligns with the proposal made by Crowston and Fagnot (2008) of this further participation as 'meta-participation', favoured by a smaller proportion of participants.

Nevertheless, there is evidence that social interaction has a positive motivational effect in terms of long-term participation. Jennett et al. (2013) note that discussion and interaction have a valuable role in driving the initial decision to participate in VCS projects, but that – particularly in gaming contexts – its true value is in long term participation. This aligns with the aforementioned feedback model suggesting that in the absence of opportunities to learn and to become part of a community, volunteers will eventually become disinterested and leave a project (Jennett et al., 2016). This is further supported by Jackson et al. (2014), with participants identifying feedback and discussion with project scientists and fellow contributors as a key factor in shaping their decision to contribute long-term to the Planet Hunters project. Furthermore, evidence from volunteer engagement patterns suggests that discussion interaction plays a role in supporting and encouraging task participation – for example, volunteers who engage most in social interaction within gamified projects display longer play sessions and contribute far more overall than other players (Tinati et al., 2014) (see section 4.8.3 below).

4.6.8 Feedback and Learning

Learning and receiving feedback from fellow players, moderators and project scientists were commonly identified motivational factors, appearing in 10 publications across a range of projects as a reason to engage in both task and meta-elements of VCS initiatives. In the survey of Zooniverse contributors conducted by Cox et al. (2015a), education-based factors received the highest overall mean score as a group, with approximately 5.5 out of 7, reflecting participants' views of the Zooniverse platform as an opportunity to learn about science in a new, fresh manner through hands-on activities. EyeWire players ranked 'learning' as the fourth most significant factor encouraging participation according to 10% of respondents, after contributing to science and intrinsic factors such as interests in science and the project being fun. Similar results were also expressed by volunteers from the CosmoQuest, Galaxy Zoo and FoldIt projects (Gugliucci et al., 2014; Raddick et al., 2010; Heaton and Torres, 2015).

It appears that motives related to feedback and learning are particularly important in leading contributors to engage in discussion facilities. Through both participant observations and interview findings, Curtis (2015a) identified learning - about the project, science and the underlying task - as key factors leading to social and team engagement within the FoldIt platform and similar findings were stated by Mugar et al. (2014) in the Planet Hunters and Seafloor Explorer projects. A later ethnographic and interview driven study by Jackson et al. (2016b) further supports this finding, demonstrating that the minority of participants who engage with discussion features shortly after joining a project predominantly do so to enhance their own knowledge and skills.

4.6.9 Technology and Web-Specific Motives

In spite of the importance of technology and the web in facilitating VCS, there is little discussion of the impact of technology on participant motivations within the sampled literature. Reed et al. (2013) and Wald et al. (2015) both identify technological issues - predominantly ease of use - as key in ensuring volunteer retention within VCS patterns, but it is unclear to what extent these factors are driven by the tools themselves and to what extent this is a result of participants' own experience in using web-based tools, or potentially, capability with the microtask regardless of technological involvement. As Baruch et al. (2016) identify however, when delivered poorly or incorrectly- for example when there is lack of use of up-to-date technology, or the services are difficult to use - there is a negative relationship between technology and participants' desires to engage with VCS projects (Baruch et al., 2016). At the same time, an interesting alternative perspective is offered by Ponti et al. (2018) with regard to technology usage within FoldIt. Highly technologically proficient players are able to create 'recipes' which use computer programs and code to automate the task process, generating large volumes of points with what is perceived to be minimal effort. This is perceived negatively by other participants, who view such behaviour as tantamount to cheating - not only because of the minimal human engagement in the folding solution, but also because of the inegalitarian nature of the solution and the lack of sharing such recipes with the wider community. In essence, then, it is the unsocial nature of such solutions, rather than the solutions themselves that enrages these participants.

4.6.10 Reputation and Recognition

Descriptions of 'reputation' within the sampled publications largely correspond to one of four related but unique categories: recognition from fellow participants, recognition from project scientists, recognition from friends and family in the 'real world' and recognition in terms of a reward mechanism, through the earning of points, badges or a reputation score. Not all of the articles distinguished between these types of recognition and so it is not possible to identify with certainty if participants perceive these forms of recognition differently, or if one form is more appealing than another. Nov et al. (2014) identified reputation-based 'reward' functions as mildly positively correlated with intrinsic motivation, but found no correlation between reputation and contribution quantity, suggesting that this increased intrinsic motivation is of no value in terms of participant productivity. On the contrary, participants perceive reputation mechanisms to be of relatively low appeal, with a mean score of 3.7 out of 7, far below the 5.98 assigned to intrinsic motivations (Nov et al., 2011a).

The desire or even need for positive reinforcement and recognition of contributions was a commonly described factor within the motivational and design literature. Participants in Old Weather enjoyed the opportunity to receive some form of validation that their contributions were both correct and useful to the science team and in the absence of such feedback or evidence of scientific progression, would become demotivated and leave the project (Eveleigh et al., 2013, 2014). Both Darch (2017) and Woodcock et al. (2017) described participants becoming demotivated by perceived failings within the data - for example, glitches, camera errors or simulated 'fake' images - and requiring reassurance by the science team that their contributions were to be used for science, although Woodcock et al. (2017) note that only a third of interview participants were aware of projects' scientific outputs. This recognition can also stem from other participants, particularly in close-knit communities such as FoldIt, where participants form groups or teams and motivate each other to push to achieve higher scores and do more work (Curtis, 2015a). Nevertheless, where community-based reputation is seen as difficult to achieve or as rewarding undesirable behaviour, such mechanisms become demotivating for low- and mid-performing participants (Eveleigh et al., 2013; Ponti et al., 2018). While the desire for recognition may drive significant numbers of participants to projects, it is a poor factor for motivating long-term engagement - in the days following the public announcement of the Hanny's Voorwerp[3] discovery, there was a notable peak of 9,400 new participants to the Galaxy Zoo project, yet 70% of participants contributed for a single day or less (Morais et al., 2015).

4.7 Engagement and Activity

A common theme within the literature is the relatively distinctive form that activity and participation in VCS projects take, with a highly skewed distribution of effort as a small and highly active proportion of the community performs more contributions than the majority put together (Spiers et al., 2018). This

[3]The discovery of a Quaser Light echo led by forum participant Hanny Van Arkel in collaboration with project scientists - see Lintott et al. (2009) for more details

participatory model is not unique per-se, in that similar phenomena and engagement patterns have been observed in related initiatives such as Volunteer Geographic Information projects and in online communities more broadly [4] (Haklay, 2016; Ridings and Wasko, 2010). In truth, however, this contribution pattern is distinct from other forms of crowdsourcing, particularly paid crowdsourcing and likely arises from the combination of the large crowd of volunteers required to complete project tasks and the volunteer basis on which many participants contribute (Mao et al., 2013a). As an example of this highly skewed pattern of participation, Galaxy Zoo I volunteers contributed 80 million classifications, yet 1,220,067 of these came from just one single participant while two additional volunteers contributed in excess of half a million classifications each (Morais et al., 2013). In contrast, the largest group of volunteers - measuring 45.35% - contributed between 10 and 100 classifications each. A similar phenomenon is seen in terms of the period for which volunteers are active within a project and in terms of the time that participants contribute for (Morais et al., 2013; Sauermann and Franzoni, 2015; Tinati et al., 2015a).

With regard to **community size**, the number of participants varies between different projects, even within the same platform (Sauermann and Franzoni, 2015). As of 2018, the Zooniverse platform for example now has over 1.4 million volunteers, although the specific number of volunteers who have engaged with any one project is far less than that and highly dependent on the individual projects themselves (Watson and Floridi, 2018). In practice, however, the *active* community is a much smaller proportion of the total user base, with a small highly active core of participants who contribute the largest volume of the workload being observed across Zooniverse projects and within highly gamified projects such as EyeWire (Luczak-Roesch et al., 2014; Ponciano and Brasileiro, 2015; Tinati et al., 2015a). This is particularly the case within FoldIt: despite having "many thousands" of registered users, participant observations have demonstrated that only between 200 and 300 of these users are actually active and in truth, the core of active participants numbers just 20 to 30 individuals (Curtis, 2015a).

Further skew can be seen in terms of the number and length of the sessions for which volunteers contribute (Segal et al., 2015). Participants who demon-

[4]The "Pareto Principle" dictates that in any community, 20% of the group will perform 80% of the work (Juran, 1954). While these specific figures do not hold entirely true for online communities, the underlying concept regarding a highly skewed distribution of effort is nonetheless commonly held to be true (Ridings and Wasko, 2010).

strate long-term participation patterns unsurprisingly tend to make the greatest numbers of contributions (Sauermann and Franzoni, 2015), although as users' active lifespans within a project increase, Ponciano and Brasileiro (2015) found that contribution patterns and levels became increasingly sporadic until eventually these users cease contributing altogether. Sauermann and Franzoni (2015) concurred with this conclusion and suggested that while long-term participants are crucial to the success of VCS projects, care must be taken to ensure that adequate numbers of new users are recruited to overcome the inevitable loss of long-term active participants. The participation levels of new participants are nevertheless highly variable and somewhat unpredictable. A large proportion of players contribute only for single sessions and then leave projects with no intention to return (Crall et al., 2017; Morais et al., 2013). Sessions can also be extremely brief, with a significant proportion of participants contributing for as little as 90 seconds and making few contributions to the project (Jackson et al., 2016b; Tinati et al., 2015b). The exact reasons for this are unknown, but anecdotal findings from the literature suggest participants are more likely to leave if their initial assets do not contain anything of note, or if they are expected to engage in tutorials, practice or simulated tasks rather than being permitted to access 'real' tasks immediately (Curtis, 2015a; Darch, 2017; Mao et al., 2013a).

Although the analyses identified within the theme of engagement are project specific, these findings were common between each of the VCS projects and platforms analysed and moreover, extend to similar communities and platforms - notably Aristeidou et al. (2017) identified similar skewed contribution levels and active periods in a weather-based CS community combining offline and online models of participation. Indeed, within the sampled literature such a phenomenon has been demonstrated and observed within at least 63 Zooniverse projects, Happy Match, Forgotten Island, FoldIt, Stardust@Home and EyeWire (Curtis, 2015a; Nov et al., 2011b; Prestopnik et al., 2017; Spiers et al., 2018; Tinati et al., 2015a). While it may be impossible to consider every possible form that VCS initiatives may take, there is adequate evidence to suggest these contribution patterns are a universal phenomenon that naturally arises from VCS based on the motivations that drive volunteers and the design and implementation of VCS projects. In summary, then, any VCS project will predominantly rely on participants own intrinsic and other motivations and this will in turn drive skewed distributions of effort as the majority of participants do not make the shift from initial motivations to the motivations such as altruism that drive long-term, highly active patterns of participation.

4.8 Sociality and Interaction

A further set of publications considered the nature of sociality within VCS and the benefits that this can bring to designers and participants. This section represents the smallest cluster of publications, with a focus on different forms of interaction within VCS, motivations and engagement patterns across social aspects and the impacts of these patterns on task activity.

4.8.1 Forms of Interaction

Discussion features within VCS projects themselves take a number of forms: blog or news pages (Heaton and Torres, 2015), custom discussion platforms such as Zooniverse's *Talk* service (Woodcock et al., 2017), asynchronous message board forum services (Greenhill et al., 2014), integrated synchronous instant messenger services (Iacovides et al., 2013), external synchronous messenger services such as IRC[5] (Curtis, 2018b) or even volunteer-led collaborative Wikis (Tinati et al., 2017a). Features are not always present from the launch of the project - a discussion board forum was introduced to Galaxy Zoo after significant volumes of emails from participants and this was later replaced by the custom Talk system due to issues with scale and also identifying new messages within the previous forum system (Tinati et al., 2015b). The extent to which these features are widespread is reasonably ambiguous. Curtis (2015a) describes such features as widespread and relatively common, present in "most" projects, but provides no citation or evidence for this statement and no specific figures or statistics regarding the frequency of such features are offered within any of the sampled publications.

In its simplest form, interaction takes the form of discussion, where participants socialise, share advice and discuss a variety of topics - both on-topic and off-topic - through the discussion platforms associated with participants (Curtis, 2015a). These discussion topics appear to be highly project-dependent, with the focus of discussion in EyeWire's chat being described as highly off-topic, while in Zooniverse projects, topic discussions are highly specialised and task-,

[5] *Internet Relay Chat*, an externally hosted synchronous chat form, where small-scale communities interact in discrete channels - each with a specific focus and community (Werry, 1996)

interface- or science-related (Luczak-Roesch et al., 2014; Tinati et al., 2017a). Participants use projects to ask questions, share their knowledge, teach, learn, build relationships or share their opinions about the project or scientists' decisions (Darch, 2017; Jackson et al., 2014; Jennett et al., 2013).

However, not all uses of the chat interface are interactive. EyeWire, for example, features a number of chat commands which participants use to assist with and enhance the task completion process (Tinati et al., 2015a). In fact, one of the most commonly used commands is *silence*, which mutes the chat window so that participants can contribute unimpeded by messages from the community. While these commands are not interactive in the strictest sense, they are nonetheless heavily used by the most active players to expand their game experience and to facilitate social interaction, particularly at the end of the completion of a task (Tinati et al., 2015a). While there is no reason to believe that commands are unique to EyeWire, there are no other publications that demonstrate the use or availability of similar commands and tools in other projects, possibly because in other projects IM-chat or other discussion platforms are not integrated within the task interface or workflow to the same extent as EyeWire.

Alternatively, projects may offer collaborative task workflows, allowing participants to contribute in groups or to build upon the solutions offered by fellow participants and group members (Bauer and Popovic, 2017; Curtis, 2015a). Within the sampled literature, the only evidence of a project in which such activity occurs was FoldIt, which allows participants to form teams and to improve upon the puzzle answers offered by other team members, with some degree of collaboration through the integrated IRC-chat system (Curtis, 2015a). There is no indication that such collaboration may occur in other projects, even in projects which are in similar domains and which share the difficulty associated with FoldIt, such as EyeWire, Phylo and EteRNA (Curtis, 2014; Tinati et al., 2015a). There are a subset of VCS games known as collaborative-tagging games, based on the ESPGame[6], which use a specific workflow where two participants attempt to tag an image at the same time in a semi-collaborative manner, but since players cannot interact with one another or see each other's responses, these are not collaborative to the same extent or in the same manner

[6] A game-based activity where participants are presented with images at random and must select a word or phrase to describe that image, attempting to guess the same word as their partner, while avoiding specific 'taboo' words (Robertson et al., 2009)

(Kaufman et al., 2016). Contrary to these projects, Mugar et al. (2014) suggest that - at least in the case of Planet Hunters and Seafloor Explorer - VCS workflows are specifically tailored to prevent this form of collaboration, with participants prevented from seeing each other's responses at any point in the task workflow.

Interaction between participants and science team members may occur either in a one-sided manner through the use of announcements and news, or through direct discussion and interaction. There is little direct, conclusive evidence within the literature demonstrating the extent to which direct interaction occurs. For example, Curtis (2018b) notes that "some" scientists find fulfilment in engaging with the community and in taking part in the community itself, while others find such activities to be "challenging." While Curtis explains that this is partially influenced by the project task, she does not offer statistics or examples from specific projects, contexts or scientists, beyond stating that this has been a core element of FoldIt since the launch of the project.

Interaction may also be external, particularly in the case of interaction between project scientists and external audiences, at conferences and through press releases (Curtis, 2018b). Jennett et al. (2013) describes more personal interactions between participants and friends, family or acquaintances who have knowledge of or a connection to the field or research topic which is the focus of a given project, stating that this connection can lead participants to engage with VCS projects and heighten intrinsic motivations. Such external interaction may also occur through external web-based or physical materials such as research papers, blog posts, virtual activities such as interactive calls or through social media, either led by scientists or project participants (Crall et al., 2017; Jennett et al., 2013; Tinati et al., 2015b). In rare cases, interaction may even take place through real-world face-to-face interactions (Cappa et al., 2016), although there is not strong evidence for the frequency or occurrence of such interactions and widespread face-to-face interaction is fundamentally incompatible with the crowdsourced and international nature of VCS.

4.8.2 Motivation for and Engagement with Discussion

Unlike task interaction, there has been much less consideration of the factors which encourage participants to engage with discussion and social features, although there is some evidence within the literature – predominantly drawn from interviews and case studies, rather than larger surveys.

Both Curtis (2018c) and Mugar et al. (2014) describe *legitimate peripheral participation*[7] (LPP) in social features attached to VCS projects as a significant factor in driving newcomers to discussion platforms. In essence, participants are driven to social features and to discussion to benefit from the knowledge that more experienced and more knowledgeable participants are able to share. In the case of Planet Hunters and Seafloor Explorer as described by Mugar et al., participants can only engage in legitimate peripheral participation through talk, as the nature of the task workflow in both projects is such that there is no opportunity to observe and learn from others' task submissions. Curtis, however, points to the occurrence of LPP in online communities within the similar domain of distributed computing, where tasks are automatic, but where participants are still driven to learn through chat systems. It should be highlighted that in Seafloor Explorer at least, Luczak-Roesch et al. (2014) found evidence of learning within the Talk community, as participants' use of terminology became increasingly specialised overtime, with initially common terms like *fish* and *thing* gradually replaced with more specialised terms such as *bryozoan* and *cerianthid*, although no significant effect occurred in Planet Hunters. This effect does not only influence newcomers and less experienced users, but also drives long-term participants and more active forum users to further engage with the community, as an opportunity to teach and share their knowledge (Jackson et al., 2014; Tinati et al., 2017b). In turn, these findings suggest an interaction between community and discussion features and participants' intrinsic motivations, as the feedback loop suggested by Jennett et al. (2016) allows for discussion-based learning to lead to increased interest in science and task.

A similar finding is the use of chat functions as an alternative to restrictive task workflows. Focusing again on the Planet Hunters platform, Jackson et al. (2014) carried out interviews with three highly active Planet Hunters Talk participants and identified autonomy and opportunities to make use of new and varied skills as a driver for Talk engagement. However, unlike LPP and learning, this motivational factor takes longer to arise, with one participant waiting until his 19th session and 279th classification to use Talk, as an opportunity to exercise his skills and contribute to subjects about which he was particularly know-

[7] "The process by which newcomers become part of a community of practice...[where] the meaning of learning is configured through the process of becoming a full participant in a sociocultural practice." (Lave and Wenger, 1999).

ledgeable or in which he had particular interest. Jackson et al., also describe a kind of branching out of Talk activities, where participants initially only observe and leave chat comments, but over time turn to other resources and mechanisms such as building collections of images or reading and commenting on project blogs. Despite this, findings relating to this effect were ambiguous, being based on comments from just one participant who admitted to only dabbling with collections for a brief time before abandoning use of the feature. This aligns with observations of the use of different features in another Zooniverse project known as Gravity Spy, where engagement with collections and other Talk features was much lower than Talk commenting and observation behaviour and slowly decreased overtime (Jackson et al., 2017).

Discussion features also serve as an opportunity for participants to self-organise and to potentially contribute more effectively or efficiently. Tinati et al. (2017b) surveyed 1,365 EyeWire users regarding their use of the real-time chat function and examined 53,090 chat messages. Managing and supporting task, team and game processes was one of the three main factors identified by participants as motivating their ongoing use of the chat system. Moreover, in terms of chat messages, the longest and most active conversations tended to be those surrounding self-organising and managing task and game processes, with such conversations spanning hours - sometimes even overnight. Tinati et al. associate these conversations with players' intrinsic motivations, noting that engaging in chat does not earn players the points and rewards associated with task completion and contrast the occurrence of these long, active, task-related discussions with discussions in Zooniverse's *Talk*. While participants *do* in fact self-organise within Zooniverse projects through the use of Talk, both Darch (2017) and Greenhill et al. (2014) describe this self-organisation as predominantly for fun and it is not possible to state based on the sampled literature to what extent such task-based organisation occurs. On the other hand, unlike EyeWire's chat, *Talk* is asynchronous and is also marked by much higher response times than other question answering and discussion platforms (Luczak-Roesch et al., 2014). Any task-based organisation would therefore be hindered by the platform itself.

As well as being beneficial, however, discussion features can lead to unintentional negative behaviours. Darch (2017) and Morais et al. (2013) describe participants sharing negative views about the research conducted through Galaxy Zoo – for example, computer science research within a project which is predominantly focused on astrophysics. In the cases highlighted by Darch (2017), the forum allowed these participants to become particularly influential and it is suggested that participants who would otherwise have been satisfied with the project became disillusioned with the project after exposure to these views. Nonetheless, these factors in turn make discussion platforms beneficial for project scientists, who are able to gauge opinions and issues within the community and also to filter issues and significant questions raised by the community, rather than using solutions such as email, which scale poorly for large communities (Tinati et al., 2015b).

Lurking behaviour appears to be common within *Talk*. Jackson et al. (2017) demonstrate that in the Gravity Spy project, Talk is one of the most popular feature in terms of views, throughout the life of the project, but discussion frequency is significantly lower across all four sub-boards: *chat*, *help*, *notes* and *science*. Similar behaviour is common in EyeWire, although the nature of the integrated chat interface means that the high proportion of players who do not directly use chat are lurkers by default and viewing behaviour is likely unintentional or non-existent among this group (Tinati et al., 2015a). Active engagement in discussion features is demonstrated to be low across projects and restricted to a small, highly active portion of the community, but it is unknown to what extent lurking behaviour is common within other projects and platform types.

4.8.3 Relation to Task

With the exception of the previously outlined experiments with virtual peers, there is no evidence for the impact of competitions and task-related sociality in participation. In terms of discussion-based sociality, while five studies present findings related to the relationship between talk and task, these findings are limited and somewhat contradictory and there is no consideration of temporal aspects of productivity.

While 90.8% of participants in projects studied by Luczak-Roesch et al. (2014) had completed at least one task, just 40.5% had made at least one talk comment. As with task contributions, talk activity in Zooniverse's Talk adheres to a power law distribution with a small proportion of participants responsible for the vast majority of activity, with highly active task participants tending to also be highly active in talk as well (Luczak-Roesch et al., 2014). However, Tinati et al. (2014) found only very weak correlation between the number of task contributions and talk comments made by individual volunteers. This is supported by Jackson et al. (2016b), who identified that users corresponded to distinct profiles, with heavy talk users contributing a moderate number of classifications and heavy task contributors in turn contributing few talk comments. However, in the case of Jackson et al, these results are based on results from just one project and only consider participants who began contributing during the course of the research period, resulting in a relatively low sample size of just 277 talk participants. Based on anecdotal interviews with participants in the same project, Jackson et al. (2014) suggest a relationship between the decision to use Talk and long-term project participation, reflecting findings from the FoldIt project (Curtis, 2015a, 2018b). Neither study offers concrete data to support this suggestion. In the EyeWire project, Tinati et al. (2015a) find that players who engage in the integrated chat message complete more tasks on average than those who don't, with a small group of highly active players responsible for the vast majority of talk and task activity. Chat participants are also shown to contribute for longer periods of time than non-chat participants. On the other hand, chat participants spend far longer using chat than completing tasks and it is suggested that chat activity predominantly takes place outside of – rather than during – tasks.

4.9 Discussion – Characterising VCS Engagement

Participants generally rated sociality and social features low in terms of the impact on their decision to contribute to a project. Nevertheless, there are a number of results which suggest an interaction between community features and participant motivation. Intrinsic motivations have been suggested by Jennett et al. (2016) to become strengthened through learning – a key motivational affordance in itself – and Mugar et al. (2014) and Luczak-Roesch et al. (2014) both suggest that this phenomenon occurs predominantly through discussion rather than task features. While altruism is less associated with sociality, there is a level of ambiguity within the literature. Example statements of altruistic motivation are generally given as a desire to "contribute to science" or to scientific research[8], rather than to contribute to the greater good. In turn, Darch (2017) suggest that participants require reinforcement and feedback that their contributions are of value for scientific research purposes, and this feedback is

[8]see for example, Raddick et al. (2013)

predominantly supplied through forums and other discussion features. In this way, although expressed by participants in altruistic terms, motivations are likely in fact motivated by some level of extrinsic motivation or collectivism, with participants willing to contribute their time and effort on the condition that these efforts result in research publications. More broadly, collectivism and community norms were also stated by participants, but there is insufficient evidence within the literature to identify to what extent this reflects explicit visibility of the resulting community or in fact, to what extent the collective efforts of the community and the ability to negotiate and discuss norms are identifiable without discussion features. Certainly, there are contradictory results within the literature, with collective motives leading to a low click-through rate for participant recruitment purposes, despite their significance as reported by Stardust@Home volunteers (Nov et al., 2014).

At the same time, there are methodological weaknesses in the motivational studies identified through the literature review process. While response rates for surveys tend to be low, the rates within the sampled literature are particularly low given the relatively large communities associated with VCS activities. Moreover, the surveys were generally conducted for just a short period of time – generally speaking just a few weeks – which given the engagement patterns common to VCS projects suggests that the majority of participants are unlikely to be aware that a survey took place. In particular, this skews results to long-term participants and thus the importance of altruism as outlined is likely over-exaggerated, given that altruism is associated with long-term contribution patterns and in turn, motivations associated with dabbling will be under-represented. There are also significant contradictions within some publications. For example, Baruch et al. (2016) found that the desire to contribute to science was a significant factor in motivating participation, yet also that the majority of participants who ceased to contribute did so because they found the assets to be of low quality, raising the question of whether the desire to contribute was truly that strong in the first place. It is ultimately impossible to identify whether survey respondents were sourced from the same group and if not, what factors initially drove dabblers to participate in projects. This, however, is not necessarily a methodological failing, given that even in explicitly studying dabblers, Eveleigh et al. (2014) found little evidence for specific factors motivating such behaviour.

It is also unclear to what extent VCS projects are designed around these motivations. While the sampled literature presented a relatively small number of methods which may increase contribution levels within VCS projects, there is insufficient evidence to understand how widespread these approaches are. Moreover, these findings are of course based on a somewhat outdated selection of projects, given that many of the projects had been retired and only relaunched at a relatively recent time. It is necessary, therefore, to conduct a project review to better understand the common features used within VCS projects to drive and encourage contributions.

4.10 Summary and Conclusions

In this chapter I have presented the findings of a literature review conducted over five databases, drawing on a total of sixty one publications. The key findings of this review are as follows:

- VCS projects are contributory initiatives where volunteers predominantly analyse data, through activities such as categorising, cataloguing, mapping or transcribing.
- Research surrounding VCS has broadly focused on four design challenges: ensuring accuracy, maximising contributions, recruiting participants and retaining these participants long-term.
- Volunteers in VCS are predominantly motivated by altruism, intrinsic interests in the science or tasks and less commonly, extrinsic motivations such as rewards.
- VCS engagement is highly asymmetric, with a small portion of players contributing the vast majority of the work and many players contributing for brief, single sessions.
- Increasingly, projects are implementing discussion platforms such as blogs, forums or custom discussion systems. These offer collaboration and opportunities for social interaction, while potentially encouraging long-term activity and increased productivity.

What is less clear from the literature is the variety of features that are common to VCS projects, particularly in terms of community and social action. In the next chapter, I will build upon these results using a review of 48 projects to understand how well these literature findings are adhered to in practice.

Chapter 5

Online Community Features Within VCS

As demonstrated by the previous chapter, it is unclear to what extent Virtual Citizen Science is a social, community-driven process. The nature of the tasks identified, the reliance on redundancy and repetition for quality assurance and the apparent lack of collaborative and more extreme forms of Citizen Science suggests an individual-driven, solitary process. Yet at the same time, community and social features appear to be associated with increased levels of participation and may be essential for learning and for engagement with VCS.

In this chapter, I will examine the presence of online community features within VCS projects, as illustrated in figure 7. Drawing on recommendations and findings from literature in the area of online communities, I will develop a framework of four themes which are associated with attracting participation to virtual community spaces and activities such as VCS. Observing the affordances and mechanisms which correspond to these themes within a variety of VCS projects, I will explore the extent to which these projects are social, community activities. Moreover, I will explore the duality of microtask and discussion activity and the predicted impact on player activity.

The rest of this chapter is structured as follows: firstly, I describe the online communities literature and findings which were used as a basis for the development of the four themes explored within this chapter. I then present the results of the project review process, divided according to theme: task visibility, goals, feedback and rewards. This is followed by a discussion highlighting key findings and their relevance to the research questions which this thesis addresses. Finally,

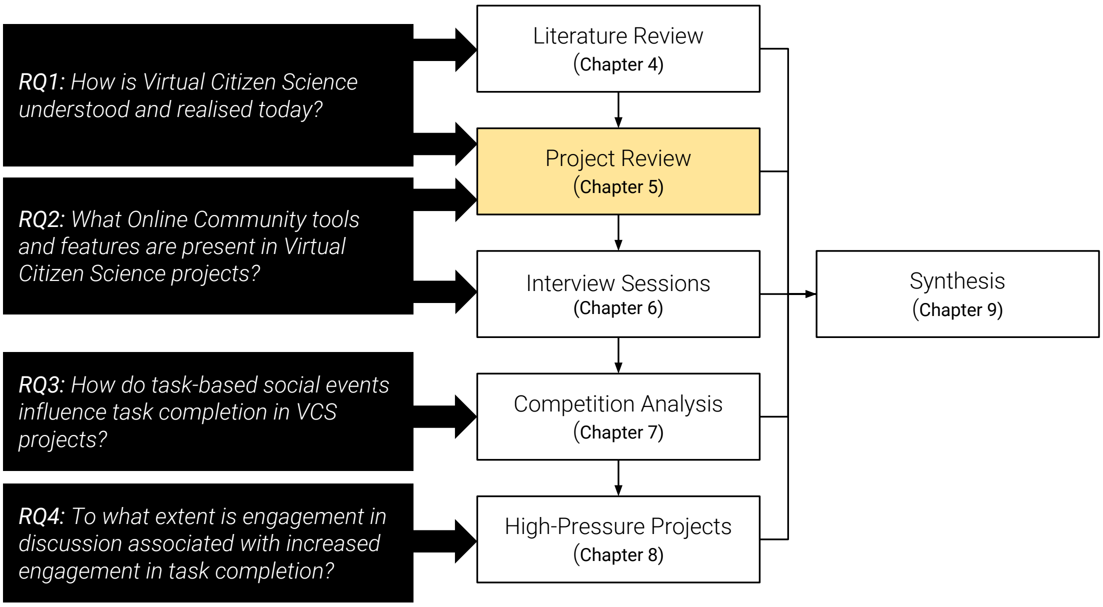


Figure 7: This second study builds on the literature review to further identify the nature of – and any online community features in – VCS projects.

I present concluding remarks and summarise the findings and situate them in light of the findings of the previous chapter and the subsequent analysis to be conducted.

A full list of the forty eight projects used for this analysis can be found in appendix D. Since the project sampling and survey process was completed, the EyeWire platform has introduced *Panoptes*: a Citizen Science design and implementation service, which allows any interested party to create their own Zooniverse project based on their own specification with the help of templates and a simple browser-based tool, which assists participants in making both front-end and back-end decisions (Bowyer et al., 2015; Vershbow, 2015). The aim of Panoptes is to significantly simplify the Citizen Science implementation process and since its introduction there has been a significant increase in the number of Zooniverse projects (Kosmala et al., 2016; Shuttleworth, 2017). However, because of the nature of the Panoptes framework, particularly its template-based nature and the relatively simplistic approach[1], projects generated through the tool are generally speaking more homogeneous and less tailor-made than those generated beforehand. For this reason, no further sampling and surveying took place, as it would be difficult to distinguish intentional design decisions from those made

[1]Panoptes was specifically designed with the intention to remove the need for project scientists to be able to code and so each project is generated using the same browser interface (Bowyer et al., 2015).

solely on the basis of the tools and recommendations of the browser tool and lack of alternatives.

5.1 Task Visibility

Echoing the findings of chapter 4, the project survey demonstrates that the microtask portion of projects offers little opportunity for interaction and sharing, nor for autonomy. Task visibility features and the frequency of these features are shown in table 7. In 44 of the 48 projects, assets were assigned automatically to volunteers, with few opportunities for volunteers to select an asset or task. The remaining four projects were either highly gamified – as in the case of FoldIt and EteRNA – or combined data collection and data analysis processes in the case of Bug Guide and iSpot Nature. EyeWire and Verb Corner also feature task/asset selection, but this is a privilege reserved for certain high-performing players (see *Rewards*). Five of these projects featured an area drawing attention to assets requiring contributions from participants, due to either time limitations, or as a quality assurance mechanism.

Mechanism	# of projects
Notification of most recent chat activity	48
Free selection of discussion threads	47
Automatic Assignment of Assets	44
Sticky/pin function	43
Entity availability limited by classifications received	41
Follow function (by thread)	34
Completion percentage (by collection)	7
Dedicated area for entities in need of input	5
View other volunteers' submissions	3
Customisable discussion feed	2
Entity availability limited by total number of entities	1
Task available for limited time	1
Follow function (by entity)	1

Table 7: Mechanisms which support task visibility.

Microtask contributions were also highly restricted, with limits placed on the availability of assets and the number of contributions which could be made.

41 projects made use of thresholds, with assets retired and no further classifications sought as assets reached a set number of classifications. One project – InstantWild – featured just twenty four assets across three categories, from automated camera traps. As new assets become available, assets are retired and so the opportunity to classify is restricted based on time and the number of assets obtained. A number of other projects indirectly featured time restrictions due to the nature of the task or subject of the research. Orchid Observers, for example, asked volunteers to gather images of orchids in bloom, which given the seasonal nature of the orchid can only be carried out during set months of the year. EteRNA featured a 'Cloud Laboratory', where participants can propose and vote on potential experiments using EteRNA infrastructure, but the large number of projects proposed has resulted in time restrictions to maximise participation. EyeWire features periodic, temporary competitions although these do not restrict access to assets – only to competition bonuses (e.g., point rewards).

Returning to the typologies set out in chapter 4, only EteRNA and FoldIt allow participants to engage in more contributory forms of Citizen Science – proposing and running their own experiments. These were not, however, contributory projects in the sense described by Wiggins and Crowston (2012), as the subjects of the experiments available to participants are restricted (Khatib et al., 2011; Lee et al., 2014). Certainly, at the time of conducting the survey process, all three of the experiments suggested within EteRNA were suggested not by volunteers but by representatives of professional scientific research bodies. Participants do not conduct the experiments themselves, nor do they engage in the analysis process after the data are generated. Even so, these represent a far greater level of participant engagement with science than the other projects described here.

In contrast to task elements, discussion platforms were much more autonomous, with almost complete freedom on the part of volunteers to select discussion threads and engage with conversations. Each of the forum services allowed participants to select and contribute to any of the available discussion threads. No retirement mechanism was present for discussion threads. Even in Zooniverse's talk, where individual retired assets have associated talk pages, talk pages are not closed or locked upon asset retirement. Each of the forum-based projects had a small number of locked threads, but these were for project-wide announcements for which no comment was required, rather than a time- or popularity-based process. In fact, discussion threads had several features to support and encourage long-term participation, such as the ability to follow or subscribe to a given thread and stickied threads to allow participants to find important threads even after many months or rarely, years. With the exception of Instant Wild, all of the project discussion systems had some form of rule system regulating behaviour and so it is not entirely accurate to describe discussion platforms as fully autonomous. Even so, each of the projects allowed for a relatively wide range of discussion topics, with each of the projects from the Zooniverse platform having a specific board reserved for off-topic discussion and general chat.

Ultimately, only three of the surveyed projects – Bug Guide, Instant Wild and iSpot Nature – allowed volunteers to view each other's submissions and this was predominantly restricted to projects with a data collection element. In both Bug Guide and iSpot Nature, volunteers can see the images submitted by other volunteers and the species identifications suggested by the community. Orchid Observers offers an interesting variant of this – volunteers are presented with the images captured by other volunteers, but cannot see the suggested species identifications, nor is there a common catalogue of image/species combinations. Finally, Instant Wild allows visitors to see the aggregated submissions for any active asset, but the restricted number of available assets also affects the visibility of these aggregations. Such restrictions on the visibility of participant submissions did not apply to discussion comments, which were open and visible to all visitors, except for the live-IM chat used in EyeWire and EteRNA and FoldIt's IRC service, which required volunteers to log-in to the system. FoldIt's IRC supports the use of channels and Curtis (2018b) suggests that individual teams use their own channels to discuss in private, but the general IRC channel is open to all who choose to participate.

5.2 Goals

While the lack of task visibility may suggest an isolated process, carried out by lone individuals who have no interaction outside of discussion tools, this is not entirely the case. Goals – particularly communal goals shared by the volunteer community – were present in a minority of projects. Most commonly, projects made use of classification challenges – a temporary drive to achieve either as many classifications as possible within the time limit, or to reach a set number of classifications. For example, *Higgs Hunters*, *Planet Hunters*, *Planet Four* and *SpaceWarps* were each the focus of separate three-day challenges carried out

Mechanism	# of projects
Classification challenges	18
Competitions	14
Opportunity for Rare Discoveries	6
Meta challenges (fundraising, attracting attention)	5
Survey (user voting for entity naming, etc.)	4

Table 8: Mechanisms which support goals.

in partnership with the BBC's *Stargazing Live* broadcasting. Each campaign challenged volunteers to – as a community – make as many classifications as possible across the three-day broadcast period. Stargazing Live campaigns have been highly successful – at its peak, the *SpaceWarps* project generated 6.5 million classifications across these three days, peaking at 2,000 classifications each second[2]. Nevertheless, while highly beneficial for the four projects, this burst of classifications is temporary. For example, two years after the challenge ceased, the challenge period still accounted for the majority of contributions to the Planet Four project[3].

In each of the 18 projects where these challenges were found, the nature of the eventual goal and the time period in which it must be reached meant that individual volunteers could not hope to achieve the goal by themselves. There were no individual-focused goals, although that is not to suggest that there were no reward strategies for participating individuals. One example of this was the *Moon Mappers* 'Million Crater Challenge', which asked the community to complete one million classifications between the 20th of April and the 5th of May 2012. Reaching major milestones – every 100,000 classifications – would result in a reward for a single member of the community who happened to make that classification. Similarly, *EyeWire* features 'Marathons' where players must complete a whole cell in less than 24 hours, receiving rewards based on the number of classifications made within this period. To emphasise the sheer number of classifications required in each case and the difficulty this poses for the community as a whole – let alone individuals – the million crater challenge was not in fact completed for some six months after the challenge period ended and the EyeWire blog features a number of examples of unsuccessful marathons.

[2] https://blog.zooniverse.org/2014/01/09/stargazing-live-the-results-are-in/
[3] http://blog.planetfour.org/2015/03/18/2-years-on-from-bbc-stargazing-live/

Moreover, while some of these challenges allowed volunteers to track the number of classifications made by other volunteers (see Rewards and Feedback), there was no change to task visibility, with no opportunity to work together on individual tasks or to view other players' submissions.

An alternate, but similar goal-driven campaign is the use of (predominantly) temporary competition events, with players driven to outperform others in terms of the number or quality of submissions made during the period. Unlike the classification challenges, which are community-based and collaborative, these competitive events rewarded individual players, or occasionally teams. In the more gamified projects, these are held regularly and in some cases underpin the gameplay itself – for example, the aim of *Phylo* is to achieve the best score through the matching task. This was also the case in Old Weather, where players move up a hidden leaderboard based on the number of classifications they have made. Players compete to achieve the rank of 'captain' and are assigned various other ranks as they progress through the leaderboard. In less gamified projects, competitions would be held almost externally to the tasks itself. That is, information about the competition and any prizes would be advertised through either the integrated project blog or via email, with little indication in the interface itself that any competition took place. In Galaxy Zoo, this included entirely external competitions: the 'Galaxy Zoo Machine Learning' competition encouraged players to develop machine learning algorithms to assist in post-data collection data analysis processes, through an external platform.

Not all competition events were linked to or rewarded the quantity of classifications. EyeWire's 'accuracy happy hour' and 'evil cubes' competitions encouraged players to be as accurate as possible. In fact, the 'evil cubes' competition is so named because players must complete 12 very difficult tasks, as selected by other members of the community. Similar to the classification challenges, none of the competition variants allowed players to view each other's submissions. Each of the gamified projects calculated rewards based on participant accuracy and so again, these events were not significantly different from the underlying task itself, but were an extension of core gameplay activities.

Conversely, not all goals involved task contributions. Five projects featured 'meta-challenges', where players were encouraged to assist with fundraising and attracting attention. Each of these meta-challenges made use of the project discussion platform either to deliver the competition – for example, to request players share information about the project on social media – or to carry out the challenge, soliciting details from players for a fundraising campaign. *Snapshot Serengeti* ran a 'Serengeti Selfies' campaign, which raised funds to sustain the project. Project scientists aimed to raise awareness about the project by publishing a book containing amusing project image assets – those in which the animals contained within the asset appeared to be posing for a self-portrait or 'selfie'. An example of this can be seen in figure 8. Players would find the 'talk' discussion page for any suitable image and add a hashtag labelling that image as a #selfie. In contrast, *Planet Hunters* solicited players' suggestions for naming new planets discovered through the project, by asking players to complete a survey or form, delivered at a separate URL. In such cases, while the *Talk* system delivered the challenge, it did not gather the player responses. Similarly, *Snapshot Serengeti* featured a second campaign – 'Save the Memes' – where players were asked to select amusing assets and turn them into 'memes'[4] (see figure 9). Volunteers created these 'memes' through an integrated button within the *talk* platform, before sharing them through social media. This campaign was highly effective, raising $36,324 and at its peak attracting 4,500 unique volunteers to the *Snapshot Serengeti* project page.

In summary, then, meta-challenges and discussion platform driven goals did not encourage players to discuss and interact with one another, but rather used the discussion platform as a medium for achieving project administration and sustainability goals. A further example of this is EyeWire's 'trivia' competition class, which players contribute to by answering questions using the IM-chat system during competition periods. Players are not rewarded for chatting and interacting with one another, but rather for being the first to give the correct answer to a given question.

5.3 Feedback

Feedback mechanisms within the literature framework are divided into two categories: task-contingent, based on how many tasks an individual or community has completed, or performance-contingent, based on the quality or accuracy of an individuals' or community's contributions (Kraut et al., 2012). Furthermore, feedback may be quantitative and systematic, through the task interface or comment/text-based, through the task interface, comments or through com-

[4] "An image, video, piece of text, etc., typically humorous in nature, that is copied and spread rapidly by Internet users, often with slight variations." Dictionary (2017)

Figure 8: Example of an asset selected through the 'Serengeti Selfies' campaign. Image Source: `https://blog.snapshotserengeti.org/2015/02/18/help-us-find-animal-selfies/`.

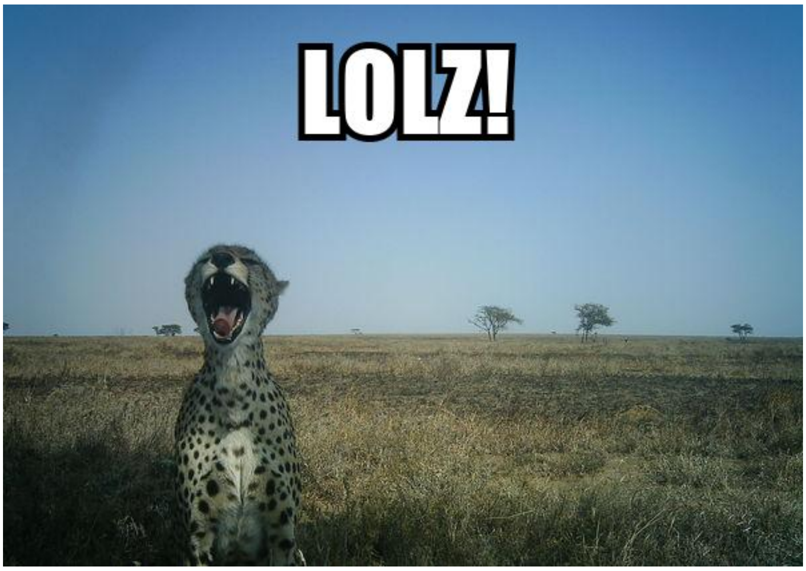


Figure 9: Example of an asset converted into a meme for the 'Save the Memes' campaign. Image Source: `https://blog.snapshotserengeti.org/2013/07/31/save-the-memes/`.

munity features such as forum discussion.

As identified through chapter 4, the effectiveness of tutorials and other training materials is questionable, due to the need for such tutorials to be skippable – particularly given the number of players who only contribute to projects for limited periods of time. In two of the projects surveyed – EyeWire and FoldIt – the training process is known to be largely insufficient to fully train players, who must gain practical experience of the task through a significant number of task submissions (Curtis, 2015a; Kim et al., 2014). Feedback, then, is potentially vital for ensuring that players learn 'correct' behaviours and to accurately complete the project task.

Nevertheless, performance-contingent feedback was relatively rare in the

Mechanism	# of projects
Task-contingent feedback	29
Performance-contingent feedback	14
Performance-contingent feedback as numerical score	11
Gold Standards for performance-contingent feedback provision	7
Progress-bars for task-contingent feedback	7
Volunteer testing	5
Majority opinion-based performance-contingent feedback	4
Comments from science team	1

Table 9: Mechanisms which provide participants with feedback on individual contributions and overall project progress.

projects studied, with just 14 projects featuring such feedback. In 11 of these projects, this feedback was provided as a numerical score, with seven overlapping with a points-based reward mechanism as discussed in section *Rewards*. One relatively unique form of performance-contingent feedback was used in the game *Phylo*; participants must pair nucleotides to build DNA sequences, which the task interprets as participants matching similarly coloured blocks – see figure 10. Participants see live feedback on the different matches they have made and these are then used to construct global leaderboards, allowing players to understand their performance relative to other players. In *EyeWire*, where players receive feedback as a point score, the score is calculated based on a number of criteria, including task difficulty and the extent to which the classification differs from the consensus received so far[5]. Since players do not receive a specific breakdown of their performance, players must instead judge their performance according to that of other players – a common factor in points-based feedback within the projects studied. This was particularly complex within EyeWire, however, as the task workflow includes additional activities such as scything – correcting cubes – allowing for the later adjustment of scores.

On the other hand, in at least one project, points-based feedback is essential for the performance of the project task. Within the FoldIt project, puzzles are very difficult and the research conducted is often complex and not necessarily presented in a manner which is accessible to players. Instead, players made use of their point scores to understand how their performance matches with what

[5] http://blog.eyewire.org/how-are-points-calculated-in-eyewire

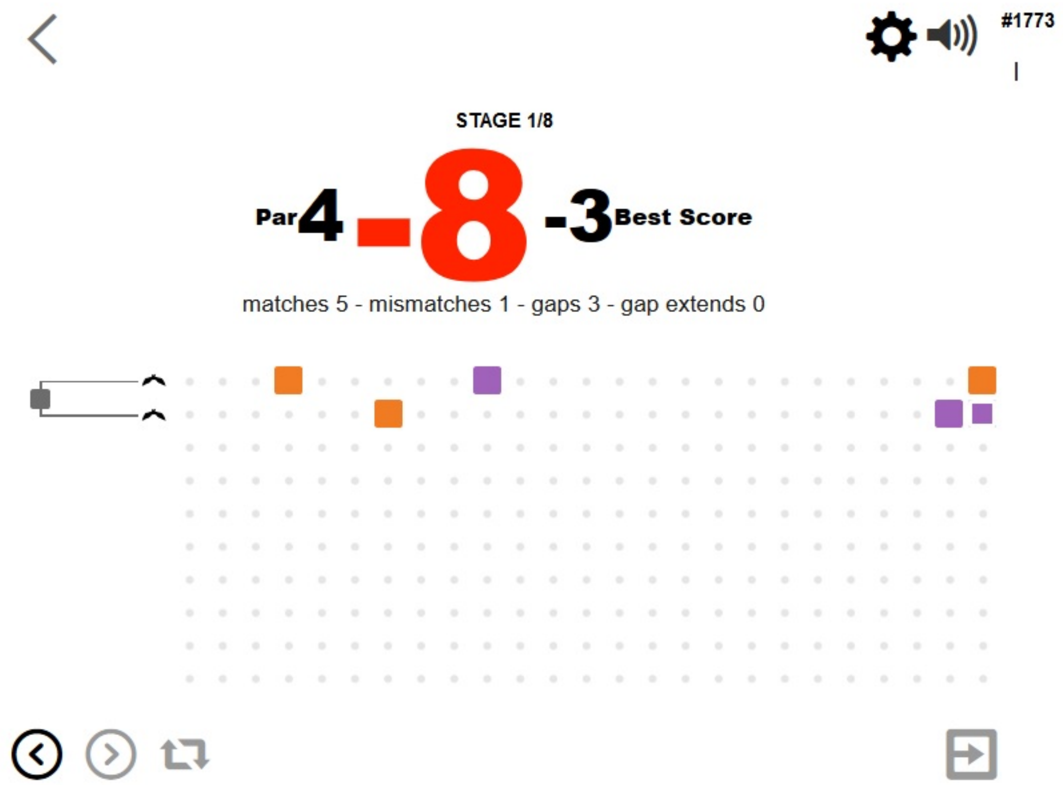


Figure 10: Phylo interface showing feedback on various stats and score, as well as nucleotide matching task as represented by coloured blocks.

is expected from them and whether they were giving a useful or correct answer (Curtis, 2015b).

The remaining four projects asked players to complete periodic tests, with a score grading performance. However, this test was largely pass or fail, with little context for the numeric score provided and no personalised feedback, or advice on how to improve the score – nor even commonly made mistakes. Stardust@Home went furthest in this area, with the explicit opportunity for players to receive feedback from members of the project science team. However, this feedback was reserved only for cases in which players had discovered a potential interstellar dust particle, which was very rare in the task itself. Furthermore, given that less than 10% of the 81 candidates had received feedback at the time at which this analysis was carried out, it is unclear if such feedback continues to be provided.

As a result, it is unclear how effective any of these performance-contingent mechanisms are in training and teaching participants about the task. It is also noteworthy that in most cases, this feedback was only available for specific assets – with seven projects using pre-selected gold standard assets for such feedback and five using tests similarly based on pre-selected assets. In itself, this lack of performance-contingent feedback is not surprising, given the difficulties with associated with ascertaining the accuracy of responses, as described in chapter 4.) However, even those projects which explicitly integrated gold standard images – for example, Planet Hunters – did not give players thorough personalised feedback, simply requesting players repeat tasks if the submission was deemed to be inaccurate. One exception to this is FoldIt, where participants modified their puzzle solutions based on the results of laboratory experiments derived from the submissions project scientists deemed to be the most suitable (Curtis, 2014).

A further four projects offered feedback based on aggregated submissions, taking the majority opinion to be correct. In the case of three of these projects, this feedback was only provided as a point score, to prevent players from knowing the exact accuracy of their submissions. Instant Wild, however, allows volunteers to see all classifications (in an anonymised list), although this is not feedback in the true sense of the word, as it is visible to any visitor to the site and is not necessarily shown upon making a submission. The literature review conducted in chapter 4 suggests that several projects use aggregated majority results to ultimately judge the overall accuracy of the project. However, in the majority of cases, these results were not explicitly shared with volunteers – either being kept anonymous (as in the case of Old Weather as reported by Roy et al. (2012)) or being shared in an academic publication (as in the case of Snapshot Serengeti as reported by Swanson et al. (2015)). However, the reason for not sharing these results with participants is unclear.

Task-contingent feedback mechanisms were more common than performance-contingent mechanisms, appearing in 29 projects. In all cases, these mechanisms tracked participants' individual contributions to a project – and in the case of Notes from Nature and Old Weather, specific asset collections – without giving additional context. Kraut et al. (2012), for example, explain that community members will benefit from fear- or benefit-based campaigns designed to inform them that their contribution has value to the community or towards the community's goal. Anecdotal evidence from VCS projects suggests that a similar effect occurs, or at the very least, that participants require positive reinforcement and to feel that their efforts matter (Cox et al., 2015a; Woodcock et al., 2017). None of the surveyed projects provided such details through the main task interface, although all of the surveyed projects did provide updates on the progress of research efforts when asked through discussion forums or through integrated blogs. Such updates will be insufficient to ensure all participants are aware of research progress, given that evidence from some Zooniverse projects has demonstrated that many participants do not view the *Talk* and blog features at all (Jackson et al., 2016b, 2017).

Conversely, by offering discussion platforms, each of the surveyed projects offered participants opportunities to give feedback to one another. The extent to which this was encouraged and facilitated by projects was somewhat variable, with *Stardust@Home* and *Herbaria@Home* featuring long-running 'stickied' forum threads with the purpose of allowing users to seek and offer feedback on their submissions. Zooniverse's *Talk* platform offers the opportunity to leave comments and captions around assets – only visible within *Talk* – as a means to share opinions and give feedback on classifications. While this appears to be a common usage, it is unclear whether participants actually receive much in the way of feedback as 90% of discussion comments sampled by Luczak-Roesch et al. (2014) received no replies. In other projects, particularly EyeWire and FoldIt which feature more complex tasks, there was no method to easily facilitate feedback among players, with the exception of feedback from promoted players known as mentors in EyeWire, who have special tools specifically for this purpose. Although in theory there is nothing to prevent science team members and adminstrators from giving feeback, only *Stardust@Home* explicitly made mention of such feedback within the project interface. The project features a specific page for potential intersteller dust candidates, each of which should – according to the page – receive a comment or comments from the science team to confirm its veracity. At the time of conducting this review, however, just 12 of the 88 candidates had received comments and there had been no new comments from the science team in over a year.

5.4 Rewards

Contemporary literature argues that rewards encourage people to provide contributions. As table 10 describes, across the surveyed projects, rewards can be supported through a variety of mechanisms. Rewards could be awarded based on the quantity of responses, as task-contingent rewards, or on the quality of responses, as performance-contingent rewards.

Mechanism	# of projects
Status rewards: Titles/Roles	41
Status rewards: Leaderboards	11
Points	11
Task-contingent rewards	11
Public announcement of achievements	7
Status rewards: Achievements/badges	5
Physical rewards	4
Unrevealed reward calculation factors	2
Privilege rewards: Additional tasks	2
Privilege rewards: Entity selection	1

Table 10: Mechanisms which reward participants and incentivise contributions.

It is important to note that few of the rewards described in the projects surveyed have any explicit value – either in the 'real world', or within the project itself. While four of the projects had physical rewards (prizes comprising of branded merchandise), these were all available only temporarily, in conjunction with an integrated competition or campaign. Other than these prizes, none of the projects featured exchangeable rewards – i.e., points, badges and leaderboard positions did not confer other rewards or prizes. On the other hand, these rewards were somewhat linked, with leaderboard positions deriving from point scores and badge rewards based on points or leaderboards.

Instead, many of the rewards derived their value from social mechanisms – either the opportunity to display achievements and rewards to other players and the wider community, or restricting rewards to a minority of the community. This restriction was particularly important when conferring titles and roles. With the exception of the Old Weather project – where players can earn titles such as 'Captain' simply by completing transcription tasks – all roles and titles featured restrictions. For example, in EyeWire, which features four roles, players must not only achieve certain criteria, but must then be selected for promotion. These promotions are only available periodically and are only given to a few players at a time.

Notably, there are two further social elements to titles and roles. Firstly, in each of the projects that use titles and roles, recipients can display these to other players, predominantly through chat features. In EyeWire, players who earn promotions gain a different coloured username in the IM-Chat interface when engaging in discussions. This is particularly notable as these colours are not otherwise indicated in the task interface – either in the leaderboard, which does not include these colours, or in the global announcements displayed to players. Similarly, in the Zooniverse and CosmoQuest platforms, eligible volunteers can display the title of 'moderator' through the discussion forum and *talk* service.

Secondly, however, these titles and roles predominantly confer access to tools and opportunities through the projects' discussion platforms. In Zooniverse and CosmoQuest, moderators are able to close and delete discussion threads, or suspend or ban discussion participants. Similarly, EyeWire's moderators police the IM-chat interface, while the mentor class is able to train new players and collaborate with them on cube completion tasks. Task-based tools and opportunities were much rarer. The only example of these tools being given to promoted players is EyeWire's Scout and Scythe player class, who are able to correct and mark incorrect cubes for repetition or further attention, as well as select assets themselves, rather than relying on the random assignment given to other players.

Beyond these rewards associated with titles or roles, access to new tasks or tools was extremely rare. Just one project offered additional tasks as a reward for participation – VerbCorner, where players earned new tasks by completing a certain number of previous tasks. Similarly, just one project restricted access to assets as a reward mechanism: EyeWire, where harder 'level two' tasks were available only to players who have completed a specific tutorial, accessible only after completing a certain number of level one cubes.

One issue posed by rewards is the possibility that players will deliberately or accidentally modify their behaviour to maximise rewards for minimum amounts of effort – in other words, the possibility of exploitation. Perhaps because of this factor, it was difficult to identify any reward mechanisms which were explicitly performance-based. EyeWire's badges include a batch reserved for high accuracy achieved through a competition class. However, apart from this, all of the rewards offered in projects were either task-contingent, or combined performance with other factors.

For example, EyeWire's points are calculated based on several dimensions, including at a minimum the accuracy of submissions and the time taken to complete the classification. However, this time-based component is complex and while taking more time increases a player's score, this portion of the score is capped at an unspecified amount. Moreover, a full explanation of point rewards is not given to players. An alternative, but similar approach was used in VerbCorner, where task occasionally awarded players with bonus points, but how and when was unclear. As a result, beyond EyeWire's specific badges, there were no solely performance-based reward mechanisms.

5.5 Discussion

There is significant variation in the level of sociality present within the sampled projects. The most social projects include the Games with a Purpose, specifically EyeWire and FoldIt, which feature integrated chat systems[6], frequent collaborative and competitive events, individual- and group-based participation and rewards and extensive opportunities for feedback relative to the performance of other players. If we consider sociality as a spectrum, then at the opposite end would be Galaxy Zoo, which features no use of collaborative or competitive goals, no reward of any kind for participants and no feedback outside of the *Talk* platform.

Moreover, in some cases there are significant barriers to entry for discussion participation. Specifically, in the eight projects with discussion board forums (except for EteRNA) players were required to register separately using distinct registration details to communicate through the forum system. This was also not a trivial process, as it took in excess of three months before my request to join the Stardust@Home forum was accepted. In addition, the structure of each project meant that accessing the discussion platforms was often completely optional and required players to leave the task interface or even, in the case of EteRNA and FoldIt, to leave the project website altogether. Given that on the whole, volunteers reported few motivations related to community and discussion, these barriers are likely to discourage engagement in discussion, even where this discussion might support players' intrinsic or altruistic motivations (such as in the case of the serendipitous discoveries highlighted by Tinati et al. (2015b)).

In each of the projects included within this review, tasks were solitary and hidden from other players, with few social or collaborative elements. In fact,

[6]FoldIt's chat system uses Internet Relay Chat which requires players to use a separate external client, but Curtis (2015a) suggests that the use of IRC is common among active players and essential for group participation and thus, chat can be seen as an integral – if not entirely integrated – part of FoldIt.

only two projects (FoldIt and EyeWire) allowed players to contribute as a team and only two projects allowed players to view each other's submissions, with one of these restricting such access to a minority of specially selected players. Even where team play was available, there was no change to basic gameplay, as participants were still expected to contribute individual solutions and could not work together to develop contributions. Autonomy was limited in every project and there were no opportunities to engage in collaborative, co-created or extreme Citizen Science through the task interface, with only EteRNA and FoldIt mentioning the chance to engage in the publication process. At the same time, discussion platforms were much more autonomous, with little structure and the chance to engage in discussions without restrictions. This suggests that discussion features may be essential for overcoming limitations in the tasks themselves. Returning to the analysis provided in chapter 4, Jackson et al. (2014) suggest that this autonomy is an important incentive for talk participation in Planet Hunters. Yet at the same time, the proportion of active talk participants in the platform is very low and Jackson et al.'s analysis draws on experiences from just three players. As a result, while discussion autonomy may in theory provide incentives to participate in and liberation from the highly restricted project tasks, in practise the extent to which this has an impact on the community is likely to be limited.

These findings raise questions about the purpose of discussion features and the reasoning behind the decision to implement such features in VCS projects. Previous analyses have suggested that – for Zooniverse at least – the talk platform is essential for task completion and offers opportunities for "peer question-answering support, serendipitous collaboration and social community building," (Tinati et al., 2015b). Conversely, this review has raised doubts about the efficacy of discussion platforms for collaboration and question-answering. The lack of task visibility renders it almost impossible for players to collaborate around tasks: response times within Zooniverse's talk platform regularly exceed several hours or even days (Luczak-Roesch et al., 2014) and the random assignment of tasks and inability to return to previous assets would make any responses moot if not received immediately. Similarly, any opportunities for serendipitous collaborative discovery are by their nature highly limited. These long response times similarly pose issues for question-answering, particularly given the tendency for players not to return to projects after completing a single, short session, as highlighted by Ponciano and Brasileiro (2015).

Viewing projects in the light of Actor Network Theory, participant activities are clearly strongly influenced by the design of VCS projects and the features and tools provided to participants. In the projects analysed within this chapter, volunteers have very little agency to select tasks and control their own activities and what agency they do have is generally a result of discussion rather than task features. To consider the example of the serendipitous discoveries that have occurred within Zooniverse projects, in which participants took a somewhat more active role in scientific discoveries, these were facilitated by talk features, rather than task activity (Cardamone et al., 2009; Tinati et al., 2015b). Discussion platforms allow participants to select their own tasks, go beyond the simple analytical processes involved in tasks and to potentially influence the research agenda (Jackson et al., 2014). As Tinati et al. (2015b) note, project scientists require assistance in filtering important queries or findings as they are simply too busy to interact with participants directly and this is reflected in the finding that just one project offered participants feedback from scientists. In this sense, then, echoing the description of non-human agency by Latour (2004), the features and technologies associated with VCS projects play a key role in controlling and preventing volunteer engagement in scientific processes.

5.5.1 Comparison with Chapter 4 Results

In contrast to the picture suggested by the literature review findings – that is, that discussion and community features are optional and used only by a minority of players – this project review suggests that community and social features play an important role in incentivising task completion. While explicit engagement in conversation may be uncommon, chat features and community-facing mechanisms lend legitimacy and value to the rewards and extrinsic motivational factors present in VCS projects. On their own, for example, badges and point scores are somewhat meaningless, particularly given the tendency for obfuscation and hidden or unclear reward schedules in the surveyed projects. Instead, these become incentives through opportunities to display badges and point scores to other players and to compare leaderboard rankings and achievements.

Interestingly, in some projects incentives were split between task and discussion elements. For example, while competition campaigns were predominantly focused on task elements only, indications of the scheduled start and end times, reward schedules and in some cases character and narrative features were reserved for discussion platforms only, or for project blogs. A good example of this is the Million Crater Challenge, which was advertised and carried out solely through the Moon Mappers blog and through the CosmoQuest forum, with no indication in the task interface itself that any such campaign was going on. Similarly, EyeWire's competitions are accompanied by blog postings which include thematic artwork, start and end dates and upcoming point rewards, none of which are indicated through the task interface or even the live IM-Chat. In fact, for some competitions, the blog also features a progressing narrative arc such as a 'Whodunnit' murder mystery, but there were no indications or reflections of this narrative arc in the competitions and interface itself – see figure 11 for an example.

As a result, it is unclear to what extent players who avoid discussion platforms and blog postings would even be aware of these campaigns. While in the case of EyeWire the competition would nevertheless remain visible and incentivise contribution through competitive activities, in the case of Moon Mappers, the effectiveness of the campaign hinged entirely on players being aware of the goal(s) and associated rewards. An analysis of the Moon Mappers blog shows that updates on the campaign did not occur and it is notable that it took six additional months to achive the million crater goal – although it is impossible to state with any certainty whether this was simply due to the goal being otherwise unattainable.

A comparison of the project review findings with the volunteer motivations identified through the literature review raises further differences – particularly in terms of rewards. While the term *rewards* implies extrinsic motivational factors, it is notable that there were no similar explicit mechanisms for intrinsic motivations and altruistic motives – the two most commonly cited motivations within the literature. For example, Segal et al. (2015) describe that targeted interventions upon classification, which clarifying the scientific purpose and progress made based on participants' contributions increase the rate at which volunteers return to projects from 6.7% to 9.7% (Segal et al., 2015). Yet such interventions were absent in the projects studied. While occasional project updates were present in all projects, either through blog or discussion systems, these appeared to be ad hoc and intended to indicate the publication of scientific results or attendance at conferences. It was also not possible in some cases to discern whether updates were official, or posted by enthusiastic or ob-

Figure 11: Thematic blog post announcing the beginning of EyeWire's murder mystery competition. Note that this 'hunt for suspects' narrative was predominantly present through the blog and was not reflected in the task itself.

servant volunteers.

Furthermore, there is a level of ambiguity in the literature with regard to the effectiveness of rewards. Few studies of motivation in Citizen Science made explicit mention of rewards, but Nov et al. (2011a) noted that participants were not enthusiastic about the use of reputation mechanics. Nevertheless, leaderboard rankings and status-based rewards were the most common within the projects surveyed. It is notable that these rewards were predominantly associated with more gamified projects and it is possible that this may explain the distinction between the literature and project findings, particularly since Nov et al. (2011a) did not explore gamified projects and Games with a Purpose. Even so, with the exception of Old Weather, all projects used full leaderboards, allowing participants to compare their performance to all other participants who had contributed to the project or collection, with some restrictions such as time or whether the contribution was made by an individual or member of the group. Unfortunately, all of the surveyed projects showed large disparities in participant performance, particularly in the case of EyeWire. This large disparity suggests that comparative feedback is not as effective as it would be were participants only shown the performance of participants with scores closer to their own, given the findings of Laut et al. (2016) and Diner et al. (2018) that comparison with high performing participants results in reduced levels of contribution and the findings of Eveleigh et al. (2013) and Kraut et al. (2012) that unachievable goals such as extremely high performing community members are in fact demotivating for participants.

The use of goals and challenges also diverged from the motivational factors described by volunteers in the literature. While collective motives were popular factors identified within the wider literature, the lack of task visibility makes it unlikely that VCS project tasks can be described as 'collective' activities. Yet even where common goals were present in projects, it is debatable whether they could be described as 'collective'. Classification challenges, for example, had very few opportunities for collective attributes, beyond indications of community-wide contribution levels. Furthermore, there were almost as many projects with competitions as with community classification challenges. Yet competition is at odds with the 'common goals' which Rotman et al. (2014) describe as being key to participant motivations.

Finally, the lack of performance-based feedback and personalised corrections is a particular point of departure between the literature and the surveyed projects. On the one hand, the importance of such feedback for project accuracy is likely to be low – particularly given the finding that projects without such feedback nevertheless achieve very high levels of accuracy through redundancy alone (Roy et al., 2012; Wiggins and He, 2016). At the same time, Jennett et al. (2016) propose a feedback model which states that in the absence of opportunities to learn, volunteers will eventually become disinterested and leave projects. It is clear that providing such feedback is inherently difficult in VCS, given the lack of methods through which a conclusively 'correct' answer to a research problem can be identified. However, a common finding among the projects studied was the lack of opportunities for learning – for example, to identify recurring errors and self-correct. This places greater importance still on the opportunity to gather feedback, where possible, through discussion platforms, given the association between discussion and learning within the literature (Luczak-Roesch et al., 2014; Rotman et al., 2014)

5.6 Summary and Conclusions

This chapter has described a review of online community features within 48 VCS projects across four themes.

- VCS projects vary considerably in terms of the features, incentives, tasks and degree of collaboration and feedback that are available to volunteers.
- Nevertheless, projects share a general lack of opportunity to view or share contributions with others, or to build on other players' solutions.
- Although projects share overarching goals, the use of specific and more tangible challenges is rare. Such challenges are generally sporadic, although more gamified projects make use of competitive events and activities.
- It is not common for participants to receive direct feedback on their submissions. Where such feedback is available, it is generally task-contingent (e.g., total submission counts) or otherwise obfuscated. Players must generally rely on discussion and player-to-player feedback plays a significant role in enabling learning and improvement.
- In spite of the importance of intrinsic motivations and altruism in VCS, rewards are relatively common across projects, conferring status or achievement indicators.

However, exactly what motivates these design decisions is not entirely clear from the literature and project reviews. To address this question, in the next chapter I present findings from a set of interviews with the team behind the VCS project EyeWire, to better understand the factors which influence the design process.

Chapter 6

Understanding the Design Process

The previous chapter discussed the large variety of online community features used within Citizen Science projects. It remains unclear, however, which factors drive the inclusion of specific mechanics within the sampled projects. While Kraut et al. (2012) associate these features with increased productivity, it is unclear whether this is true for developers and to what extent they consider these mechanisms to be appropriate for use in VCS contexts. This chapter – as illustrated in figure 12 with respect to the preceding chapters – aims to explore the processes involved in the design of VCS projects, with a particular focus on what factors influence the inclusion of specific features and mechanisms. For this purpose, I conducted three semi-structured interview sessions with six full-time members of the EyeWire design and administration team as outlined in chapter 3. I begin by outlining the context of the EyeWire project, including the task and scientific purpose. Following this, the findings of the interviews are presented, grouped into four themes which arose during the data analysis and coding process: project needs and success, observed volunteer motivations, the four online community themes form the project review framework and the role that sociality plays within EyeWire itself and in terms of productivity. Finally, I conclude with a discussion of the implications that these findings have for this thesis.

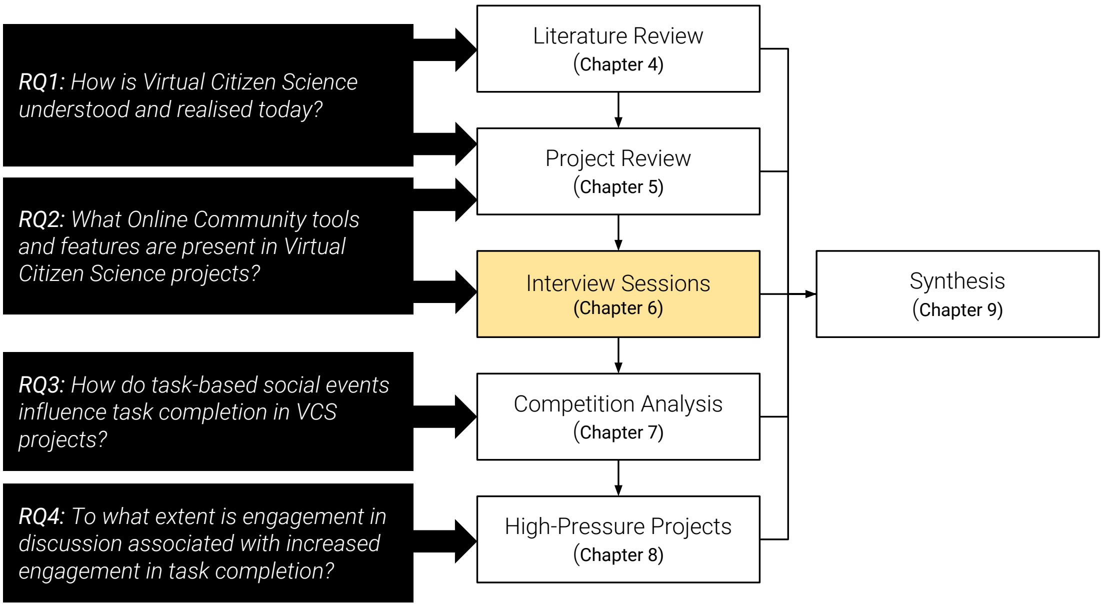


Figure 12: The interview process drew upon evidence and findings from the project review, to understand factors that influenced the implementation of online community features within EyeWire.

6.1 Project Context – EyeWire

EyeWire is a crowdsourced Citizen Science initiative with strong game elements, where participants identify neuron cells in Magnetic Resonance images of mouse retina and trace the path of these neuron cells within each image to enable future research (Kim et al., 2014; Tinati et al., 2017a). The project corresponds to the *data analysis* category identified in chapter 5, with no data collection processes and participant activities not extending to the publishing, data preparation or research question definition stages (Tinati et al., 2015b). Players are assigned a 'seed' segment of a neuron by the Eyewire algorithm and must trace the path of this neuron from the seed to the opposite side of the cube, with multiple trace submissions aggregated to ensure accuracy (Kim et al., 2014). Scan images are presented to players in the form of a three-dimensional portion of the larger neuron cell. Players contribute by mapping cells within two-dimensional cross-sections of the three-dimensional image, with many two-dimensional segments making up the overall three-dimensional 'cube' which participants submit, as illustrated in figure 13.

EyeWire is complex compared to many other VCS tasks (Kim et al., 2014). Although players do not need specific expertise to contribute to the project, the barrier for entry to the project is high, as to contribute effectively, players must develop their tracing skills (Borrett and Hughes, 2016; Lieberoth et al., 2014).

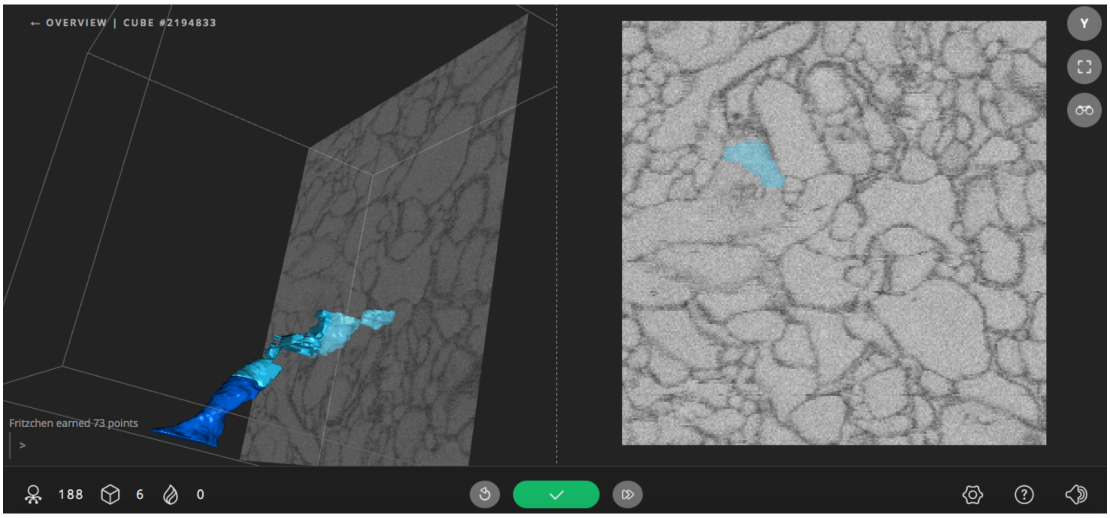


Figure 13: EyeWire tracing interface, showing left hand 3D cube view with seed and tracing, right hand 2D cross-sectional view and bottom left corner live chat overlay.

In contrast to FoldIt, which is similarly complex and features a lengthy but optional tutorial, the EyeWire tutorial is short, yet compulsory, but is insufficient to train players in how to trace, with the most successful contributors investing “tens of hours” and “thousands of cubes” while learning how to trace (Curtis, 2015b; Kim et al., 2014)

6.2 Themes and Heading Categorisation

As noted in chapter 3, a phenomenological approach was used to analyse and code the interviews and define common themes or categories which the EyeWire team's experiences share. This chapter presents those common themes as headings, with distinct subheadings in each category that arose from the commonalities and terms that the EyeWire team used, as well as from the literature review and project review findings. The aim of this chapter was specifically to understand the tools and features present in EyeWire, but also to understand what factors had influenced the introduction of these features. As a result, it was through this lens that the interviews were coded and assigned to themes.

The first of these themes is project needs and success. This theme is a central concept within the EyeWire design process and by extension, throughout the interview sessions, as it is based on these needs that strategic decisions about the project were made and the resulting project grew. Although the specific needs described by the team diverge somewhat from those identified in chapter 4, they are presented here as identified through the interviews rather than the literature, as each of the participants identified common issues and these issues were subsequently referenced by participants when explaining their experiences.

An additional theme discussed heavily by participants was *volunteer motivation* as perceived by the interview participants. It would be unreasonable to expect the team's predominantly anecdotal observations to accurately reflect volunteer motivations and it is unsurprising that the motivations described by the team do not fully or accurately map to the motivational factors outlined in chapter 4. Nevertheless, these perceptions similarly proved instrumental in the decision-making and design processes associated with EyeWire and the team's views are presented here predominantly on this basis. Moreover, it should be highlighted that in the absence of these perceptions, the decision-making process in EyeWire can otherwise seem counter-intuitive – for example, with the team prioritising competitions despite the outcome of surveys such as that conducted by Tinati et al. (2017a). Although approximately one fifth of survey respondents contributed to EyeWire for fun (20.12%), the largest proportion did so out of a desire to advance the aims of the project (28.92%) or of science generally (26.49%). Only a very small proportion of respondents mentioned competitions as a factor motivating their participation in EyeWire (1.21%) and this was a smaller proportion than those participants who played EyeWire to relax and avoided competition (6.67%). An emphasis on competitions, then, would appear to have the potential to alienate more EyeWire users than it would motivate, although it is unclear to what extent those users who find EyeWire fun do so due to the use of competitions.

Having established the basis on which the design decisions were made, the third theme identifies the specific online community features included in EyeWire, as well as further rationale for their inclusion. These design decisions are subcategorised based on the different themes used during the project review, as the interview process is intended to support the project review process by explaining why certain observations may have been made. No other broad categories were described by the EyeWire team, either individually or collectively and this is likely due to the high reliance on goal and reward strategies within the project.

Finally, the role of sociality in EyeWire is considered. This theme was highly emergent and was included both on the basis of the overarching research question explored within this thesis and due to the emphasis placed by the team on the importance of sociality in EyeWire.

6.3 Project Needs and Success

Throughout the three interview sessions, participants described three key factors which heavily influenced the design process. These factors are: ensuring overall efficiency, providing a fun and game-like activity and public dissemination. This section outlines these factors and the influence they have had on EyeWire and provides a grounding for subsequent discussion of the EyeWire design process.

6.3.1 Efficiency and Cost-Effectiveness

Ensuring EyeWire is as efficient as possible has underpinned many of the design decisions made by the team in recent years. Unlike other VCS projects, the tracing task within EyeWire was previously a full time job, with professional neuron tracers employed to complete much the same tasks that volunteers do now. As a result, the VCS task output from EyeWire is frequently compared with the work and workload performed by full time professionals:

> "The way we gauge success from a scientific standpoint is 'are we getting these cells mapped? Are they getting mapped accurately? Is it faster than paying people to do it? Than paying tracers to do it?'" (*Participant 1*)

Although volunteer submissions were compared with professional traces in terms of both quality and quantity, participants 1 and 5 noted that ultimately quantity and efficiency were the more significant concern. This was true even though participant 5 noted that volunteer traces were no more accurate than the traces provided by professionals:

> "We were trying to go faster. We were trying to complete more cells, faster, because at the end of the day, it does cost money to run EyeWire. And it has to be cheaper to do this than to just pay tracers and surprisingly, after a couple of years of Eyewire it wasn't really that much cheaper." (*Participant 1*)

> "There was a push to try and make EyeWire more efficient than it currently was. Because for a time, when you really did the maths, it was like the amount of money that was being put into paying employees to run the Citizen Science game itself was more than the money being put into paying tracers at Princeton to just do it themselves and the quality of the data was not necessarily better on our end. So it was like, what is the point?" (*Participant 5*)

In fact, while participants described the need to assess both accuracy and productivity as the two main success criteria within EyeWire, these two facets are not necessarily compatible. Throughout the sessions, participants described a number of trade-offs resulting from the desire to balance the need for accurate, yet efficient volunteer tracing. For example, participants 5 and 6 described an intended change that the team had hoped to make to the quality assurance process within EyeWire, potentially sacrificing the quality of traces to allow for an increase in quantity. This change would be accomplished by removing the need for redundancy for a certain proportion of the player base, such that high performing players would be weighted accordingly and allowed to trace cubes without the need for aggregation. Ultimately this feature had to be scrapped for mathematical issues with the underlying algorithm:

> "It was gonna be crazy. It was gonna mean that some people who were really good were going to be able to do cells almost single-handedly, in that their probability of being right would be so high, that the system would trust them enough to spawn new cubes from what they were doing, for instance." (*Participant 6*)

This drive for efficiency plays a role in restricting the features which can be added to the project. In the majority of cases, participants pointed to a lack of time and resources as the main reason for an inability to make desired change to EyeWire:

> "I feel like most of our features that we want that haven't been implemented are based on time and priorities that they haven't gotten built... There's a variety of things, but I'd say 90% of the time it comes down to we just don't have the bandwidth to do everything." (*Participant 4*)

This lack of resources not only prevents the introduction of new features, but also has left the team unable to overhaul existing features that are unfit for purpose. In particular, the game masters and community director all felt that the tutorial process was – and had long been – insufficient to teach participants how to trace cubes effectively, but the need to prioritise project efficiency not only influences the features that can be added, but also the time available to the team to make changes to the platform:

> "Our tutorial could use an overhaul." (*Participant 4*)

> "Tracking how many people, like, get through our [tutorial process]... I don't remember those percentages off the top of my head except that like they're all predictably low." (*Participant 6*)

While this drive for efficiency was predominantly financial, it also stemmed from the needs of the project scientists. Participant 3 introduced a further trade-off between accuracy and productivity – how the needs of project scientists compare to the needs of participants:

> "The scientists at Princeton say we need to get more work done and we either need to make sure that the meshes are more accurate or more focused so sort of our features can either be driven by the community or the scientists which is interesting." (*Participant 3*)

Where these two clash, the team have had to reduce engagement and motivation among the player base by prioritising the scientific research needs of the project. This can best be seen with regard to the format of meshes[1] which are presented to participants. While the science team needed meshes to be more accurate to aid with research, the participants perceived the changes negatively:

> "We changed the way that we generated the 3D meshes. It's very new geometry. It had a different look to it... They'd definitely grown attached or were very used to the old way that things look." (*Participant 3*)

> "The interesting thing about what [participant 3] is saying is that sort of the meshes that [we] ended up developing were sort of superior to the previous meshes in basically every way. They had sort of fewer discontinuities in them and they were more scientifically accurate, but ironically the community had grown accustomed to the discontinuities in the old meshes." (*Participant 2*)

These suggestions are not unique to EyeWire. Gardiner et al. (2012) associates increases in accuracy with increased cost and reduced research pace across citizen science projects. Even so, this need to balance accuracy and productivity highlights the importance that even small gains in motivation can hold in ensuring the efficiency of projects. The trade-off and at times, incompatibility between the serious, scientific nature of the project and the need to appeal to players was a recurring theme throughout the interview process and can be seen as a notable force in shaping the design decisions made throughout the history of EyeWire.

[1]The team use the term 'meshes' to describe the cross-sectional scan in which participants click to conduct tracing – that is, the right hand square within figure 13

Figure 14: Screenshot of three different puzzle interfaces taken from the game Monument Valley, described by participant 1[2].

6.3.2 Fun and Engagement

One factor that the team felt was very important was that participants would play EyeWire for fun and perceive the task as a game. Participant 1 summed up the design process in this regard by making comparisons with the popular mobile games Angry Birds and Monument Valley (shown in figure 14):

> "We're trying to be like a real game. A game that people play for fun, that millions of people play. So we look more at like Monument Valley and Angry Birds and these more light-hearted puzzle games." (*Participant 1*)

This process, however, was fraught with challenges resulting from the need to balance fun with the serious nature of the underlying scientific processes. Both participant 3 and 5 explained that the main source of fun in EyeWire is not necessarily from a game-like activity, but rather from the opportunity to produce a 3D scan from a 2D image, or to make something from nothing:

> "One of the main things it's always come down to, especially with Eyewire, is basically you're able to click on a point and something happens. So, this kind of ability if you want to get esoteric in a sense, to create magic, right? You're able to create something by doing something." (*Participant 5*)

[2]Image Source: `http://www.funoon.co.uk`

> "I think we had a class at NYU and one of the guys basically said that the one area of fun that we have is the satisfaction of clicking in the 2D and then seeing a 3D model appear. It's a very basic interaction there." (*Participant 3*)

Participants each identified several factors which pose barriers for making EyeWire into a game. The first of these was the type of gameplay and the underlying activity. Each of the participants described a game genre with which they felt EyeWire was incompatible:

> "There's no way I could say anything about what makes a game like Eyewire or any... Community/multiplayer game successful and like, have it apply to something like Myst." (*Participant 6*)

> "I think it depends a lot on the context of the game, the genre you're shooting for, is it a fast game? Is it a slow game? Are we solving puzzles? Are we mutilating aliens?" (*Participant 2*)

However, the largest barrier to participants 5 and 6 was the absence of progression and a final goal to achieve, a feature which they felt was central to the experience within many games:

> "Certainly if you want to say that the ultimate reward systems, in some games, is to give them the feeling of having beaten something... Almost no citizen science game is going to give that feeling. Unless the outcome is 'you cured cancer! Good job!'" (*Participant 6*)

> "I think one thing that maybe is important, for some games at least and it feeds into our system as well, is that even if you are doing fairly repetitive tasks, it's nice to know you're getting better at them... Even though there isn't a level system, even though you aren't playing level 20 because [in EyeWire] there is no level 20." (*Participant 5*)

As suggested by participant 6, however, the issue is not that the game lacks a narrative to overcome, but rather how this ties into the feeling of being rewarded. All participants felt that the most significant element of any game-based system is the provision of rewards for player activity:

> "There's some stuff that's universal, like you have to get adequate rewards for what you're doing." (*Participant 5*)

> "It needs to be not just that you're getting rewards, but it feels like it's proportionally appropriate?... I don't know, if Mario were going along just collecting coins and not ever getting any coin stuff, it would sort of be like 'woo, he's getting coins? Who cares! The coins don't...translate into anything!" (*Participant 6*)

Even so, it is important to consider that EyeWire is very much a rewards-based game. EteRNA, Phylo and also the target games described by participant 1 such as Monument Valley (shown in figure 14) make use of very puzzle-based interfaces, where there is a clear goal to reach and the experience derives from solving the puzzle. This is not the case in EyeWire – participants earn points as a reward for tracing, but the interface itself is not a game in the conventional sense, beyond the interaction described of being able to make changes within the interface:

> "Your core task in Eyewire is map a branch of a neuron from one side of the cube to the other and the reward that you get for doing that cube is you get points." (*Participant 1*)

There are further tensions between the nature of VCS and the design of EyeWire as a game. Participant 2 stated that the team felt that on the whole, VCS games were unengaging, while participant 1 did not feel they were real games:

> "Well, one thing that we've kind of said in the past that our goal is not to be the best citizen science game, because that's maybe not to high of a mark." (*Participant 2*)

> "In our eyes we're not trying to be like a citizen science game. We're trying to be like a real game." (*Participant 1*)

Still, the team did not point to specific examples of citizen science games that they felt were not real games. On the contrary, participant 2 pointed to EteRNA as having precisely the look and feel of a real game:

> "I think there are some games, ETERNA, for instance, that feels like a game. It literally feels like a great game on the internet – it's engaging. The graphics are cool. The interaction is really fun." (*Participant 2*)

Moreover, whatever the team's perceptions of citizen science games, the nature of EyeWire as a citizen science activity leads to a number of restrictions in terms of what the team can add to the project, due predominantly to funding:

> "There are people who're like serious gamers who have like, requests and then we go – Well, we're a citizen science game (laughs) we're not a full game studio so I think that... Knowing our limits is key in some of our feature development." (*Participant 5*)

> "Most citizen science projects have much lower budgets, right? I expect you would see much less bells and whistles. You won't see all the things that are really expensive to create unless you're a really well funded citizen science project or unless you're really adept at getting people to help you build things." (*Participant 1*)

In addition, the serious nature of the game means that the team have to think carefully about the kinds of activities and imagery which are featured within the project:

> "I would say, I mean, there are a lot of games that are violent out there and I wouldn't expect a citizen science project to be a violent game. In the inverse, citizen science projects are usually building, you know, they're more like world builder games, rather than destruction games... And though a citizen science project is competitive, it's not as competitive. It's like, collaborative and competitive at the same time." (*Participant 1*)

> "EyeWire as a brand has a voice in the way that we speak to our players and we have expectations that the players expect us to behave in a certain way... So we might come up with a competition that is pirate themed, which is sort of classic past or we might come up with one that is space pirate themed." (*Participant 2*)

Ensuring a game is fun, then, is a difficult task within a VCS space. Projects are associated with significantly lower levels of funding than other games and there are features that are simply incompatible with the serious nature of virtual citizen science activities. Even so, at their core, projects can be engaging and share an activity that – no matter how *esoteric* it may be – captures the imagination of participants and can function as a game even if the activity remains relatively repetitive and work-like.

6.3.3 Public Dissemination and Awareness

A lesser, but still significant element influencing the design of EyeWire involves public dissemination and spreading awareness of the project, neuroscience and science more broadly. Participant 2 noted that the team explicitly monitor social media to confirm whether participants are sharing news from the project:

> "I think that one other component in addition to how much work we're getting done and user engagement, is how much traffic are we getting on social media? Are people sharing our images, are people sharing their achievements in the game? How are we being talked about? I think that's another really interesting metric." (*Participant 2*)

On the other hand, participant 1 was not sure to what extent EyeWire as a project includes options to share activity to social media platforms:

> "I think right now, there's not a lot of easy ways to tweet and share stuff. Basically when you sign up for a competition, there's an option to Facebook it or Tweet it... Or when you get a badge you can Facebook it or Tweet it. But those are really the only two times that we like, explicitly invite players to share things." (*Participant 1*)

Participant 5 extended the importance of public awareness more broadly, from social media to general knowledge or awareness of the project among the public more generally:

> "I mean, there's other ways to gauge success, too, for a project, is, you know, there's always the virality of it – the popularity of... Most people have heard of FoldIt or of Eterna and you know, it's... Have people heard of this, are they having people interested in it?" (*Participant 5*)

Conversely, however, participant 1 felt that one of the greatest strengths of citizen science was the opportunity to share science with a greater audience:

> "I would add there's kind of another definition of success in that we're reaching a lot of people... Allowing a lot of people to see neurons who've never seen them before. To maybe spark curiosity in people... And citizen science really showcases this." (*Participant 1*)

At the same time, however, participant 1 explicitly drew a comparison between public awareness needs and ensuring the scientific outputs of the project:

> "Yeah, there's all the other externalities of like visuals and like science communication and what not, but if our goal is to map neurons, we better damn well be going faster than paid people who are mapping neurons all day." (*Participant 1*)

Ultimately, though, no matter how successful a Virtual Citizen Science project may be in terms of public dissemination, it is by scientific outputs – by productivity and efficiency – that the project is ultimately judged. It is unclear to what extent this holds true for all projects, particularly those which have never involved the work of paid individuals on a large scale. For example, Lintott et al. (2008) describes how galaxy classification was carried out by unpaid individuals, in an unsustainable and inefficient manner. Yet, it appears that efficiency and productivity underpin perceptions of the importance of Citizen Science initiatives. To return to the categories described by Wiggins and Crowston (2011), these projects are not *education* projects. The main aim of the project is not to raise awareness of issues, but to complete scientific research.

6.4 Volunteer Motivations

A prominent theme within participants' responses concerned the perceived motivations of EyeWire volunteers and the behaviours that aligned (or failed to align) with these motivations. Across the three interviews, participants described a diverse variety of motivations and motivational factors, corresponding to specific volunteer behaviours and the feedback and responses from the community to specific features and events. Participant 2 described a cross-section of motivations, aligned to different members of the community, which he felt were important to consider when designing and developing EyeWire and the associated tools and features:

> "Are they motivated by points? Are they motivated by joy? Are they motivated by the community? And I think that we have players that are motivated by all of these categories." (*Participant 2*)

6.4.1 Altruism

Despite the importance placed on altruism by the wider literature, none of the participants mentioned altruistic motivations directly. However, participant 1 described altruistic motives as part of a longer list of factors that she believes drive participation in EyeWire:

> "The people who want to play Eyewire are people who are interested in making a difference, right? They want to help us understand the brain." (*Participant 1*)

Discussing the issue in more depth, participant 1 explained that EyeWire – the game, project and community – must rely on altruism if it is to function effectively. For a brief time and prompted by external sponsorship, EyeWire offered monetary rewards to certain participants, based on performance (i.e., as competition rewards). This policy was disastrous in terms for player behaviours, prompting cheating, botting and low quality submissions, but more importantly had an extremely adverse effect on the overall community and environment, leading to reduced engagement:

> "I think when you pay people to do well, it incentivises a different type of person... It also opens up a dialogue of: why aren't we getting paid? We're doing work. You have people who work for you who are doing stuff really similar to us, they get paid. Why don't we?... It created suspicion within the community. So if someone came in and started doing really well on the leader board, the players were like: are they really playing or are they cheating? And that's not the vibe that you want." (*Participant 1*)

Conversely, throughout all three interviews, participants placed a large emphasis on the importance of rewards: tangible rewards such as merchandise and intangible rewards such as points and badges. This emphasis was supported by the participants' experiences of the launch of the Scythe Complete feature. In spite of communicating with the players to explain the importance of scythe complete for overall project accuracy and efficiency and relying on players' altruistic natures, the team were disappointed to find engagement with the feature was initially extremely low. Participants 2 and 3 described the feature as 'failed' or 'a failure' and participant 1 described the preparation for the launch in detail:

> "We did numerous beta testing sessions with the players. Showing them how to do it. And getting their feedback on Scythe complete and trying to explain to them why scythe complete was important and why we need to go and mark the branches as confirmed/complete. We built this complicated new feature into the game and then we launched it and nobody used it." (*Participant 1*)

When pressed on why they had ignored the feature, players considered it to be too much work:

> "They were like 'Oh, well, we just – We dropped down on the leader board. It just feels like work.' I think it felt like work because there was no reward... [Then] we built two entirely new sets of badges, complete with like the thing... And with that, they started using it. (*Participant 1*)

These findings echo those of Darch (2017), that participants are not motivated to complete workloads which they perceive to be unworthy of their time, although in the case of Galaxy Zoo as described by Darch, there is no reward for players to earn. This raises questions regarding the extent to which altruism truly drives participation in VCS – particularly in more gamified projects and those with rewards. Within EyeWire, at least, altruism is insufficient to drive participation in more complex or slower tasks and players expect to be rewarded accordingly for participation. This was not unique to the scythe complete feature, as participants 2 and 3 described a similar issue with the generation of meshes as described previously. Clearly, then, for the most active players, points and associated rewards are the most important factor governing participation. New features and changes, regardless of how beneficial they may be to science are perceived negatively by this highly active group and there appears to be a trade-off between overall project efficiency and individual player efficiency – or at the very least, players' perceptions of their own accuracy. This may at least partially result from the highly gamified and point-driven nature of EyeWire gameplay, but it is important to consider that the self-reported motivations from EyeWire players as described by Tinati et al. (2017a) do not strongly differ from those of other projects. Altruism is one of the three strongest reported motivational factors reported by players – yet this does not align with player behaviours and feedback.

6.4.2 Intrinsic Factors

Intrinsic motivations are highly significant in Citizen Science initiatives and EyeWire is no exception to this (Tinati et al., 2017a). Participants did not refer to intrinsic motivations at length, but nonetheless noted interests in the underlying science, neuroscience and the brain itself to be key motivational factors for EyeWire participants:

> "They want to participate in a research lab. They want to communicate with scientists. They want to interact with other gamers who also like the brain. They're a pretty brainy bunch! Like not to be punny, but they really like to talk about science and tech articles and things." (*Participant 1*)

In addition to the underlying scientific aims of the project, participants also mentioned players' interests in the technology and infrastructure that enabled the EyeWire project:

> "Though I do directly talk to the community and even about you know technology, because they're so passionate about Eyewire they also get a bit passionate about the technology of Eyewire, making sure the servers are working perfectly." (*Participant 3*)

In fact, contrast to the findings of chapter 4, participants had contradictory opinions on the importance of the underlying scientific research for EyeWire players. Participant 6 did not believe that the opportunity to be involved with the publication of research was motivating for all players, while participant 2 felt that contributing to science was a key driver for engagement:

> "So I know of a couple of papers that have published for Eyewire, players have been mentioned in the paper as authors and... So... However... That's something that's a little bit nerdy, in a sense, it's not a... It's kind of a different fulfilment for some people, than others... But that's really outside of the game, in itself. So creating some kind of system inside might need to be artificial... It would be a good motivator for people to do it. Do the science." (*Participant 6*)

> "I think that the players take that to heart. They feel like they're really contributing to the science and enjoying it in the process." (*Participant 2*)

However, while the wider literature suggests that volunteers' intrinsic interests – particularly in the underlying science – are most important for driving the decision to participate in a project, the team suggested that this was not entirely the case:

> "It's very hard to say for sure whether anyone cares or whether that curiosity or that 'huh, that's so cool looking' turns into something later down the road... One thing that we've actually kind of struggled with, with Eyewire is I think, you know, it's gotten really serious looking and intense looking and it's really important to make it fun. And engaging. And appealing to people enough to start." (*Participant 1*)

At the same time, however, participants noted that player motivations – particularly intrinsic motivations with regard to interest in science – are not always compatible with the realities of the scientific process.

> "It had been two months since the previous hangout and one of the players was like 'so what have you learnt from our neurons since last time'? And I think one of the researchers was like... 'Oh, we haven't even looked at them'. You don't do that. That was the truth, but it kind of made the player feel like crap, because they're all like 'oh man, we've been working so hard!' Doing all these neurons... So in a way, to ease the burn of the slowness." (*Participant 1*)

Returning to the theme of feedback from chapter 5, it becomes clear that while players benefit from task-contingent feedback and from feedback on scientific progress, this can have negative impacts on players' intrinsic and altruistic motivations if progress is perceived as too slow – which is often the case.

On the other hand, it appears that delaying or avoiding such feedback has less of a negative impact on player motivation and engagement and participant 1 explicitly described preventing interaction between players and scientists as a solution to this problem. Participant 6 raised similar issues, in terms of the impacts of the delayed gratification common to VCS and scientific initiatives on player engagement:

> "There's sometimes a need to have some sort of immediate duel satisfaction or... Something like that, that isn't necessarily immediately available in a science game. So, in that case, sometimes you do need to create some kind of artificial, little extra kind of thing." (*Participant 6*)

The decision to design and frame the core tracing task and wider project in terms of a game was originally driven by feedback from early tracers and students that the tracing activity held an element of fun:

> "Completing the work does have an element of fun that is why this thing even started. It was a research tool and they thought there was some element of fun there." (*Participant 3*)

However, as described previously, participants 3 and 6 stressed that this was not necessarily fun in the most conventional game sense, but in an "esoteric sense", where players enjoy having the power to manipulate the in-game world, clicking and interacting on an object to transform it. This form of fun has wider, overarching implications for the design of - and player engagement with - the EyeWire project. Expanding on the theme of fun and intrinsic motivations, a commonly recurring issue within each interview session concerned participants questioning the nature of EyeWire as a fun or at least, game-like experience. Participant 3 was particularly concerned with this issue:

> "I guess I've been trying to make the task itself a game and that's always been a big question, you know... We call it a game, but is it a game? So I've been trying to see how we can make that experience more game like." (*Participant 3*)

Participants 5 and 6 gave more detail when discussing this issue, noting that the core EyeWire task shows little variation and is thus very different from other games – even those which rely on puzzle-based game play. :

> "We've applied [features] to Eyewire as far as kind of knowing that we, unfortunately, have no variation in our task. What we do have is at least a variation in the typography, but not in the action." (*Participant 6*)

6.4.3 Social and Community Factors

In spite of the importance of 'fun' as a part of the EyeWire task, a commonly recurring theme within each interview was the perception that tracing cubes could be monotonous – particularly during longer play sessions. Participant 5 saw this as a key issue within EyeWire and thus believed a central responsibility of the game master role was to use chat to make the game more engaging:

> "You know, coming up with fun stuff to do so that people continue to be excited about this extremely repetitive task that we're having them do for hours on end." (*Participant 5*)

All 6 participants viewed the integrated chat system to be a crucial element of EyeWire - in many cases second only to the points and reward systems:

> "We also have chat, which is a big part of our community. They... They talk to each other and enjoy it as a social means, they also use it to communicate about the work that they're doing." (*Participant 4*)

Nevertheless, participants 1 and 2 went beyond chat, viewing the wider community and associated interaction as the most important factors for participant engagement. In the case of participant 2, community and interaction forms part of the central game itself and is analogous to interaction with AI opponents, such as in the case of competitions:

> "I would say one of the most important factors is community, I think community is crucial, I think creating ways for your community to engage with your project and each other, that's really important... [Players] being engaged with [other] players, whether they're humans or whether they're AI, in EyeWire's case we have both, which is pretty interesting." (*Participant 2*)

Conversely, participant 1 described community and interaction as a reward mechanism, where players are congratulated and receive positive reinforcement from their peers:

> "You're online to get your badge and to get your points but also you're congratulated by your peers because it's publicly noted in chat and there's like big golden things that pop-up in chat which you can click on to see other players' badges. And everyone's like 'congrats, way to go,' so I think it adds an element of like human reward and kind of community reward to it... You want your community to be embracing and when someone new comes in and is kicking ass, you want your community to be congratulatory." (*Participant 1*)

Expanding on this, participant 1 described a reward cycle where chat reinforcement offers players the opportunity for progression, self-improvement and eventually promotion:

> "It becomes a kind of grass-roots levelling up system so rather than in most games there's a very clear path to get to the next level, in EyeWire it's like if you're active in the game and you're talking to people in the chat, then the players might say 'hey, are you interested in being a scout?'" (*Participant 1*)

Nevertheless, it appears that – for participant 3 at least – chat and community are not inherently necessary features of EyeWire. Chapters 4 and 5 have suggested that Citizen Science activities are relatively solitary, individual-centred activities and EyeWire is no exception in this sense. Nevertheless, chat plays a crucial role in providing freedom and allowing participants to take a break from the potentially monotonous and repetitive core activity:

> "Our normal cube experience could almost be done – if it was fun – could be done without a community. It's independent. It's not required to communicate." (*Participant 3*)

Participant 6 likewise described a group of players known as 'lurkers', who are characterised by low levels of chat participation and mid-low levels of task participation. It is unclear whether this low level of task participation is a result of a lack of integration into the overall community, but participants 4, 5 and 6 noted that this group were relatively unresponsive to community activities (particularly competitions):

> "However, if we talk about that weird grey middle ground of people who aren't power players, or who aren't what we would call newbies. They aren't as vocal in chat per se." (*Participant 6*)

> "They didn't seem to do any more than they usually do." (*Participant 5*)

In summary, chat is a secondary feature of EyeWire: one that is neither required to carry out the core tracing task, nor commonly used by the entire community. Nevertheless, this feature is viewed as highly significant by the design team and one they have observed to be highly engaging for the more active portion of the player base – a portion who respond positively to the relationships they build, the positivity offered by fellow players and the freedom and escape offered by chat participation.

6.4.4 Extrinsic Factors and Rewards

As previously described with regard to altruism, both participants 2 and 3 expressed the belief that it is the reward schedule of EyeWire more than any other factor that influences volunteer behaviour:

> "Some of our top players, they play until it kind of no longer is a game and it's a way of life. You know, they put in 40 hours a week, so... It's about pure efficiency, so they'll turn off the transitions between tasks as they call them, cubes as our other players know them, because they just want to get a high score, higher score all the time..." (*Participant 2*)

> "[Players are] driven by the points." (*Participant 3*)

This heavy emphasis on points and extrinsic factors clashes with the intrinsic and altruistic motivational factors identified in chapter 4 and the team were all too aware of this. In fact, participant 1 identified implicit contradictions in the statements made by participants in feedback to the team (and other external individuals) and their in-game behaviours.

> "They like to say that they don't play for points, but they do." (*Participant 1*)

In spite of the lack of consideration given to extrinsic factors throughout the surveyed literature, the team described a key player behaviour associated with points and reward scheduling. Players would prioritise tasks according to the volume of points available for completion of a task and how quickly they could accumulate points through a given activity:

> "Anything that contributes towards your points score is definitely more incentivised than things that don't." (*Participant 4*)

> "I think people like the points score the most, not just because it's on the leaderboard, because it's kind of the core interaction that they're familiar with, but also... It's a value that can accrue the fastest in a sense. Like, you can get more than one point at a time, for a cube. Where as every other stat in your profile derives in a simple one-to-one ratio of some kind... They're usually more excited about their point score, because it's usually going to be their biggest score." (*Participant 5*)

Nevertheless, it should be noted that EyeWire is inherently a points-based game, as described by the team themselves. It is not necessarily surprising that players are motivated predominantly by points, as this is the core interaction of EyeWire. In addition, as the team noted, many of their views are influenced predominantly by ‘power players’ – long term and highly active players, who by their nature are those who are driven by the inherent points-based gaming activities. Players for whom rewards-based gaming is of less interest are less likely to remain with EyeWire or at least, are less likely to engage heavily with such points-driven tasks. In fact, participants 4, 5 and 6 touched on the presence of such less engaged players who they characterised as lurkers, noting that such players were less responsive to reward mechanisms:

> “Now there were some newer players, who were maybe very interested [in rewards] and they were getting really involved in Eyewire, initially and some of those players did want to [win prizes]. And that’s because they were excited. However, if we talk about that weird grey middle ground of people who aren’t power players, or who aren’t what we would call newbies... They aren’t as vocal in chat per-se... We do have lurkers. They were less likely to be excited to win something like that.” (*Participant 6*)

> “Yeah, I think they weren’t like turned off that this was happening, but they didn’t seem to do any more than they usually do.” (*Participant 5*)

In addition to reduced engagement in features such as scythe complete, this prioritisation also raises issues with productivity and efficiency. While the most important factor in determining efficiency is the number of cubes traced, overall efficiency is also dependent on other activities such as scything, which do not offer such large point rewards. Moreover, this focus on rewards has had unintended consequences among the top performing players. EyeWire determines accuracy based on a majority voting algorithm, with administrative actions (scything and completing) performed after cubes are gathered, to correct the aggregated cube and player scores. As a result, more accurate players can initially earn a lower score, before corrected cube submissions are made.

> “Their scores are based on consensus so it’s based on what other people have determined to be right, rather than a hard and fast, you know ‘this is the right answer’. So I think that can be hard especially for new players, just starting out, to not know immediately about whether they did something right or wrong.” (*Participant 4*)

This phenomenon persists in EyeWire even now and – when combined with the importance players place on points – has led to difficulties in encouraging the completion of simpler cubes, as players are unwilling to take a hit to their accuracy and resulting score – no matter how temporary this may be:

> "There is some issues with accuracy being shown, in that certain higher ranked players prefer not to play on easier cells, because they're more likely to be compared with lower ranked players. And that can temporarily affect their accuracy." (*Participant 6*)

> "Oh yeah. Even temporarily – They're so sensitive about it that like, 'guys don't worry, things will get reaped and fixed, your accuracy will go back up, don't worry'! They're like 'no!' Slipping out of the 95% rank, it's not amazing any more." (*Participant 5*)

6.5 Online Community Framework

During the interview process, participants made comments which aligned with each of the online community themes within the project review framework, as well as specific features that were observed in EyeWire and additional projects. This section outlines those comments and what they suggest about the design process of VCS projects as online communities.

6.5.1 Task Visibility

Although Kraut et al. (2012) suggest that task visibility can motivate participants to do more work, the EyeWire team have progressively restricted the option to choose individual tasks. This was predominantly due to the observation that newcomers would join the project, complete a tutorial and then begin work on a very difficult cube, resulting in players becoming demoralised and leaving the project:

> "We realised that sometimes we're dumping people who have just done 10 tutorial cubes into a very advanced cell. And it's no wonder that they were really confused and leaving immediately. So we differentiated the task based on difficulty." (*Participant 1*)

As with all of the projects surveyed in chapter 5, EyeWire has no opportunity for group or team participation. However, unlike the other projects – where this lack of team participation appears to be a result of the reliance on majority voting to produce aggregated responses – the team showed a keenness to implement team-based participation. The restriction, as identified by participants 1 and 3, was not a result of underlying accuracy or scientific concerns, but due to a lack of resources and difficulties posed by the core gameplay of EyeWire and lack of specialised roles that could support team-based play:

> "One thing that we've talked about a lot but it's a little tricky to figure out exactly how the interaction is gonna work is kind of to get people to work together on a team, um... I know we've talked about different roles, like everybody kind of specialises in a different area. However, EyeWire is kind of like this one... It's a fairly narrow skill set. Well we have a few different areas with maybe a slight difference in skill is required. But there's not so much specialisation. It's just kind of like how good you are." (*Participant 3*)

> "We've built a notification system, a pop-up system, a better infrastructure for joining teams. An idea for how players could form their own teams and compete. There's so many things that we have imagined [but not implemented]" (*Participant 1*)

Although players cannot view each others' submissions, EyeWire features an implicit form of task visibility. The project overview features an overview of currently active cells, with completed cube segments coloured according to the specific player(s) or more broadly, specific player classes (see figure 15). The main value of this feature, however, is not in its use by players, but the opportunity for the team to see at a glance how work within the project is progressing:

> "I'm looking at a cell right now, that cell on that TV screen is all purple. Dark purple. That means that the players have all scythe completed it and if the game masters go to all the cells and they all look dark purple, that means the players are using scythe complete and we don't need to run stats in order to see that." (*Participant 1*)

Participants 4, 5 and 6 discussed a different form of task visibility, concerning the narratives and artwork associated with competitions and in-game events. While all participants felt that these were important features for engaging and motivating a portion of the player base and also a significant responsibility of the game masters' role, participant 5 in particular noted that the team are largely unsure if players actually view and experience these blog-based features:

> "The people who are slightly less dedicated, but still like, play a lot, I think tend to more just come to the game and not look at our blog, not look at our Facebook page, stuff like that. So they wouldn't necessarily always know." (*Participant 5*)

While participants viewed VCS as collaborative, they had clearly struggled to introduce true collaboration to EyeWire. The restrictions in task visibility were not the result of intentional decisions designed to prevent players from interacting, but were instead necessary due to issues of engagement and accuracy

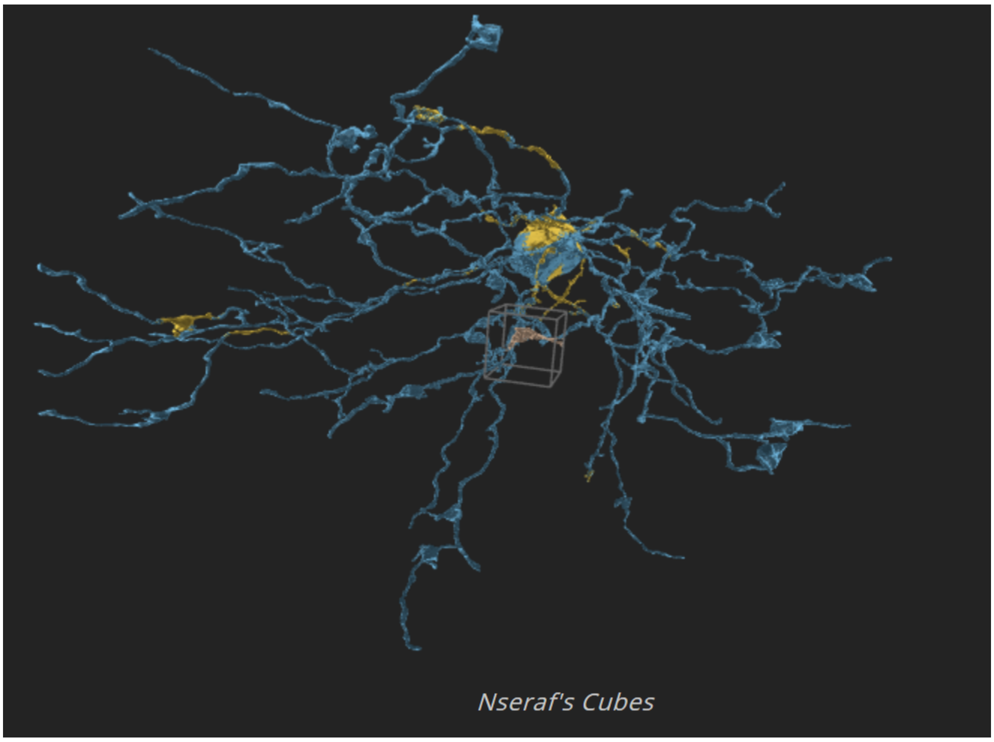


Figure 15: EyeWire cube visualisation interface, showing a specific player's completed segments (highlighted) of the currently active cell (darker segments)

within the project. Despite the team's willingness to introduce team-based play to EyeWire, the realities of VCS and lack of diversity make this difficult. VCS tasks are designed to be completed by all players, while the team viewed task-based activities as requiring different players to engage in different roles. This suggests that a truly, team-oriented VCS activity may currently be difficult to implement. On the other hand, Crowston et al. (2018) notes that Galaxy Zoo Quench failed to achieve its goals due to a lack of coordination, suggesting that as projects move towards more collaborative citizen science models, teams and roles will become increasingly vital for research.

6.5.2 Goals

The team were motivated to introduce goals to EyeWire partially in response to weaknesses in the inherent nature of VCS. On the one hand, VCS projects have explicit goals for participants, which are a core strength of the project, particularly in terms of volunteer motivations, something which participant 1 was particularly aware of:

> "[Our advisors] from like Warner Brothers and Blizzard and stuff – what

> they've said is that we have something that they would like, never have. That they would kill to have, which is purpose. And that is something that Citizen Science projects should play up. You know? I think no matter how gamified that you become, you shouldn't lose your purpose" (*Participant 1*)

Even so, participant 6 called the effectiveness of these goals into question, noting that the scientific research goals of projects are simply too high level to be appealing or accessible to the player base:

> "Well, I guess kind of like you said – the problem with most citizen science games is you can't win... Unless the outcome is 'you cured cancer. Good job.'" (*Participant 6*)

This was something participant 1 also reflected on, noting that "science goes slow." As a result, participant 1 felt it was important for any game, but particularly EyeWire, to offer opportunities for the player base to 'win' while playing:

> "It should make you feel like, you know, you're winning along the way. Right? Every little success in a game – It's giving your brain that anticipated reward, that dopamine release that makes you feel good when you're playing." (*Participant 1*)

Goals were also motivated by a rapid fall in traffic observed after the launch of the game. After an initial period of high traffic and contribution rates, submissions to the project quickly began to fall – a factor which Sauermann and Franzoni (2015) suggest is common to VCS projects:

> "We launched Eyewire, there was a huge bump in traffic and then after like, a month it just dwindled back down and we were like 'crap, what's going on?'" (*Participant 1*)

Participant 1 explicitly felt that this was linked to a lack of goals, or more specifically a lack of competitions. Player 1 did not expand in significant detail why this might be the case, but instead stated her opinion that competition is an integral element of any game. This conclusion is somewhat surprising given the comparison made to puzzle-driven mobile games such as Angry Birds and Monument Valley by the same participant:

> "You know, a game without competition is arguably not a game." (*Participant 1*)

In response to this drop of traffic and the assumption that competitions were a significant game feature that the project was lacking, the team decided to trial an event to encourage team-based competition. This did not allow players to collaborate, but instead teams pooled points based on the social network from which they had heard about EyeWire:

> "We realised that most of our new traffic was from social. So it was like Facebook vs Reddit vs Google Plus... Vs Twitter vs Team X – the veterans, the players who'd been playing since public beta." (*Participant 1*)

Competition infrastructure was missing from EyeWire at the time of the event and this led to significant issues which had not been foreseen prior to the event, with a significant portion of the workload lost due to a merger – an impossible seed which is spawned by the algorithm in response to previous errors. Even so, while the competition was not entirely successful in terms of getting work done, it was still highly effective in driving participants to EyeWire:

> "It's probably because we don't have any competitions! These people expect a game. Let's give them a game... So what we did is we called it the Eyewire Games and it was a battle of social networks... Actually unfortunately during that competition there were some server issues and there was a huge merger that was grown and because we didn't have a way to remove them, Team Reddit got hit and a huge chunk of the work that they did was removed... But I think it turned out okay and it rejuvenated our traffic and it set the precedent for 'okay, maybe the site is going to do ad-hoc competitions, maybe we should keep an eye on it'. And our traffic kind of slowly was rising." (*Participant 1*)

Since then, the team have introduced regular competition events, held fortnightly, monthly and quarterly. Participant 1 went on to describe that this is due to the effect that these competitions have on player numbers and on overall productivity:

> "The themes for the competition, the narratives that go with them, the art, the timing of the events, these are all coordinated and created by our game masters and our illustrators. But they do prove very effective in getting players to come back and to up their participation." (*Participant 1*)

As well as productivity, however, there are also competitions which tie into the desire for accurate tracing, as described by participant 6:

> “We have lots of different competition formats. We don’t just have periodic ’get all the points you can’ sort of things... Like it’s competitive, but it also trains their skills, like Evil Cubes... Every week we usually have a Happy Hour, which is just, like, giving special bonuses based on how many points you get in the first place... Our big competitions, we also do an accuracy Happy Hour, which is trying to do the same thing, but also be as accuracy as you can in the process, because then you’ll get special accuracy-based awards as well...” (*Participant 6*)

Yet an additional beneficial outcome of competitions is interaction and engagement within the community. Even so, it should be highlighted that the team specifically felt the need to introduce a competition format for this purpose, suggesting that competitions alone are insufficient for community engagement, no matter how social they may be:

> “And we also have our trivia, which happens during those competitions that has absolutely nothing to do with game play itself.” (*Participant 6*)

> “It’s community engagement.” (*Participant 4*)

Trivia is not the only source of community engagement, however. The team observed emergent behaviours among the player-base in response to competition outcomes, with a positive chat atmosphere and many taking the opportunity to congratulate each other on their achievements. This has increasingly resulted in the use of a second, smaller event at the end of competition periods, serving as an “awards ceremony”, where players are encouraged to share their achievements with one another:

> “This award ceremony at the end of Happy Hour is where we declare all the winners, where all the bonus points are won... Because you get badges for being a winner in these categories as well. And everybody kind of comes on line and it kind of takes 45 minutes to go through all the events and give all the awards and chat is just like blowing up and everyone’s super excited... You’re online to get your badge and to get your points but also you’re congratulated by your peers because it’s publicly noted in chat and there’s like big golden things that pop-up in chat which you can click on to see other players’ badges. And everyone’s like ‘congrats, way to go’, so I think it adds an element of like human reward and kind of community reward to it.” (*Participant 1*)

Ultimately, then, goals – tied to competitions – have become a central, recurring feature in EyeWire, tying into the core gameplay but also driving participant retention and productivity. Beyond this, however, goals also encourage interaction and sociality among the community which may have a further impact in motivating and encouraging participation within the projects.

6.5.3 Feedback

Participant 1 felt that one of the significant changes made to EyeWire was in ensuring players felt that their time spent within EyeWire is appreciated:

> "If you click on your profile... you can see all their stats, like how many cubes they've done... How many trailblazes they've done, so how many times they've been like the first player to do like a cube, so... That may not be a game mechanic, it's more an element of the interface, you know the data we expose to the players in order to allow them to highlight the time and effort that they've put into Eyewire." (*Participant 1*)

In the early stages of EyeWire, there was little opportunity for players to view their own performance in terms of accuracy. Performance-contingent feedback was one of the most common requests made by players, who did not understand if they were doing well at EyeWire or if they were performing poorly:

> "There was a leaderboard but you couldn't see – You had no idea how well you were doing. You couldn't see your accuracy. There was a period of time where players – they would email us, asking for their accuracy. Because they had no idea if they were like sucking at doing Eyewire or if they were excelling at it. There was no ranking. Every player was the same." (*Participant 1*)

Even so, while EyeWire offers points, leaderboard positions and congratulatory messages to players based on performance, participants viewed these entirely in terms of rewards. Participants complete cubes and they earn points and leaderboard positions for this effort:

> "The reward that you get for doing that cube is you get points" (*Participant 1*)

> "[The main game mechanic] would be... they receive points and that goes up on the leaderboard." (*Participant 5*)

This extended to participants' perceptions of how players viewed task-contingent feedback. Since these indicators increase one by one, participants felt that points were more valuable to players because point counts increase far faster:

> "I'd say that those categories can be less effective than the regular points, because they don't currently appear on the leaderboard or get contributed towards your regular score." (*Participant 4*)

Even so, participants did feel that performance-contingent feedback played an important role in EyeWire. Since opportunities for progression are generally lacking, participants 5 and 6 felt that the accuracy bars displayed on player profiles served as an opportunity for continued self-improvement and indicator of progression, to overcome the lack of a levelling or narrative function:

> "But there's also their accuracy bars and for the players that care about the quality of the work that they're doing, which I'd say is fundamentally the majority, the accuracy bars are pretty crucial to their incentive, as well." (*Participant 5*)

> "Yeah, a lot of games do have a levelling system. So in a sense by having an accuracy bar, that kind of..." (*Participant 5*)

> "...is the levelling system" (*Participant 6*)

While feedback is clearly important for players, who see it as an indicator of their achievement and their progression within the game, the team noted issues with exposing such data to participants. The decision to not provide performance-contingent feedback in many of the earlier projects may then not be a result of the difficulty in calculating and providing such feedback, but a conscious decision to avoid negative consequences from allowing participants to see how they are performing.

6.5.4 Rewards

Throughout each interview, participants placed a great deal of emphasis on the importance of reward mechanisms – be they points, badges or physical prizes – for player engagement within EyeWire. As noted when discussing extrinsic motivational factors, EyeWire is very much a reward-driven game and participants were very clear that points and associated rewards are central aspects of the core gameplay loop. All 6 participants named points and badges as the most important game feature within EyeWire:

> "The most important? Probably, you know, points is up there. With the most important, because you know, your core task in Eyewire is map a branch of a neuron from one side of the cube to the other and the reward that you get for doing that cube is you get points." (*Participant 1*)

> "What makes it a game rather than a job? So primarily that would be when a player commits a task, they receive points and that goes up on the leaderboard." (*Participant 6*)

Both participant 1 and 6 felt that ensuring players are adequately rewarded is a central premise of any game. While on the one hand, each stressed that games may feature other less direct reward systems (for example, the feeling of progress), the participants stressed that any game – apparently regardless of genre – should offer some form of reward:

> "And any game, if it's a game, generally speaking has a reward system." (*Participant 6*)

> "A game should be fun, it should be rewarding, it should be challenging. It doesn't have to be social, but fun and challenging. And it should make you feel like, you know, you're winning along the way. Right? Every little success in a game. It's giving your brain that anticipated reward." (*Participant 1*)

It is notable that this is a large point of departure between gamified and non-gamified Citizen Science experiences. As suggested by chapter 4, VCS generally relies on intrinsically rewarding activities – the feeling of having contributed to science, advanced research or learnt about specific fields, to name a few. Extrinsic factors and rewards were even suggested to have a negative impact on volunteer motivation. Yet from the perspective of the EyeWire team, the project is first and foremost a game and for this reason, must offer rewards that are fair, have some level of context and retain their value as players earn more:

> "An important nuance to any reward system, is that it needs to be not just that you're getting rewards, but it feels like it's proportionally appropriate? Like you don't just wanna get like – I don't know, if Mario were going along just collecting coins and not ever getting any cool stuff, it would sort of be like 'woo, he's getting coins? Who cares! The coins don't give him any special powers! They don't translate into anything!'" (*Participant 6*)

A further departure from the literature concerned the practise of offering prizes for participation – a practise which, as demonstrated by chapter 5, is relatively rare in VCS. Several participants discussed their past policy of offering such prizes (described by the participants as 'swag'), originally prompted by a corporate partner. In spite of the issues described by Kraut et al. (2012), participants found that far from driving low quality submissions and poor behaviour, swag was largely beneficial and in particular, had a strong community building element that the team had not expected:

> "That was a motivation for them too, I think, so the community aspect and wanting somebody else who hadn't got the prize to, you know, get it once they had already won." (*Participant 4*)

As with other features such as social features, discussion platforms and goals, participants observed a clear distinction in behaviours and motivation between power players (along with occasionally enthusiastic newcomers) and the mid-performing lurkers. However, participant 6 felt that this was predominantly due to the loyalty and commitment these more active players had for the game, more so than with other features, where participants suggested a lack of engagement from lurkers:

> "In most cases, it seemed to be that the players who were more advanced seemed to be more interested in getting those prizes. Mainly because they seemed to have some sense of duty or loyalty to Eyewire, so they wanted to have the brand that they were you know, a player of this game." (*Participant 6*)

> "I think that it was usually the most involved power players who got the most interested in Swag." (*Participant 5*)

Overall, the team perceived rewards as highly motivating for players, although it was unclear from their statements to what extent this reflects the comments and behaviours of the players. Participant 5 drew a link between players' decisions not to leave the project and the use of reward features. In contrast, participant 1 drew more heavily on her own opinions, stating that players enjoyed rewards, but not giving any specific examples as evidence for this:

> "I think generally the players like any kind of reward that you can provide them." (*Participant 1*)

> "In our case we have a point system that, while we've tweaked it over the years, is pretty fair. I think, personally. And a lot of people seem to basically accept it, so people stay motivated to be involved because they know that what they put in is largely going to be what they get out of it as well." (*Participant 5*)

6.6 The Role of Sociality

During the interview sessions, participants made reference to the important role that sociality plays within EyeWire. This section outlines the value and use of the live chat service within EyeWire, both for players and for the team, which broadly corresponds to three themes: game management, task enabling and productivity.

6.6.1 Game Management

When asked about their processes for game management, such as avoiding bugs or negative reactions for the community, all 6 participants noted that they lack formal procedures or mechanisms for such management activities. Participants noted that they rarely observe quantitative data – not due to a lack of time or interest, but due to a lack of necessity. Instead, in the event of any issues with the game, players will often draw the game masters' attention and ensure that feedback is funnelled to the developers or community manager:

> "We don't really have a formal data-driven mechanism for that. We mostly – because we have a real time feed with the players, if something really bad happens, we know because they're in chat." (*Participant 1*)

> "If something's changing, we don't want to say 'we're going to give more work to you'. It's not like that, it's like 'we're going to give you more trust. You know, you have the ability to do things that you weren't in the past.' I don't just say that in a scoffy kind of way... Because they're not just usernames, they're people. They're real people who are enthusiastic and they care and they want to help." (*Participant 1*)

Offering more details, participants 2 and 3 felt that chat (and by extension, other player communications such as email) offered a unique window on the health of the project and community, which could not be offered by mere statistics alone:

> "[Game masters] either exchange emails with the players or just elicit feedback directly from chat usually, right?" (*Participant 2*)

> "Yeah, I think they are definitely kind of how we view kind of the health of the community." (*Participant 3*)

This process has proven particularly important when assessing the effectiveness of particular features or diagnosing issues with community health:

> "After a couple of weeks, maybe a month, they weren't really using it very much and so we just kind of started – I just started doing chat interviews with the players and I was like 'why aren't you guys using this'?" (*Participant 1*)

6.6.2 Enabling Tasks

When discussing community management, the game masters identified three tasks with a common aim of enabling task (and by association, research) completion. In particular, after the game first launched, when tutorial features were lacking and players had yet to grasp how best to trace cubes, chat, sociality and collaboration were essential for the successful running of the game:

> "Our progression has always been, at first as a community manager, it was really to talk to and just walk players through learning the game, back when we didn't have as many tools for them to learn how to do it. So there was a lot more hands-on hand-holding in a sense, training, getting those players up to speed and learning what was good, what was wrong and what they should do... In the game. And through that we managed to start having relationships with these people." (*Participant 6*)

Over time, this role has been transferred to the players, with a specific player role (*mentor*) introduced to assist with this purpose. Participant 1 noted that the team believed that discussion with other players was important to accomplish what the tutorial could not:

> "How does a player learn to be good at Eyewire?... We think it's by talking to other players. Sometimes you'll open up and a player has done 10,000 cubes and their accuracy is still 65%. That's rare, but sometimes that happens!" (*Participant 1*)

Nonetheless, participant 5 suggests that the game masters continue to engage in this process, which still remains important to EyeWire and the community as a whole – although it has in fact evolved. Leveraging these relationships and through the use of chat, game masters continue to assist players, but more significantly enable task completion through encouragement and ensuring player retention.

> "There are maybe three core tasks of community management in my head which are: help, encourage and retain. And helping can be anything from, like, helping people with language problems, to actually helping them with how to play the game, helping them get better at the game, like that sort of thing. Encourage is a particularly broad set of activities... [Including] talking in the chat and getting people pumped up about things." (*Participant 5*)

In addition, participant 3 touched on a particularly interesting benefit of sociality and chat participation which has not been touched on previously – collaborative action to enable new, more difficult forms of gameplay and task activity than players can achieve on their own. In particular, at the highest levels of play (i.e., for the most difficult cubes) individual players cannot be expected to trace, scythe and complete such cubes on their own. At the same time, game masters are too busy and themselves may not be individually able to complete such cubes accurately (to return to an earlier quote: “there are no right answers.”). For this reason, collaborative play through the chat interface is crucial for these activities:

> “The higher level of game play does require a lot of interaction with our community managers, so I guess I would say it’s... It’s in order to keep the game kind of feeling lively and then working with the top players, in order to do those more complex user interactions.” (*Participant 3*)

6.6.3 Relationship with Productivity

Participants 1 and 3 both suggested a possible relationship between chat and player productivity. Participant 3 suggested that chat can counteract the effect of having to complete multiple cubes in succession:

> “I’d have to say chat, because with a small community and they get to know each other and it kind of makes up for maybe the dryness of the default work that we’re having them do.” (*Participant 3*)

This ties into the findings of Tinati et al. (2015a) that a large proportion of players mute chat while completing cubes, but spend more time chatting between cubes than they do completing the tasks themselves. Chat is used to break up repetitive activities such as cube completion, rather than being an integral part of the task itself. Even so, the team have integrated indicators of task achievement into the chat window. These include adding colours to the chat interface which indicate high performing players, so that the best – and at times only – way for a player to demonstrate how well they have performed is through the use of the chat interface:

> “They also do receive a special chat colour, based on their status, which they can use to display, like, what level of gamer they are.” (*Participant 4*)

> "So when you become a ranked player your username changes colour in chat and since, right now, the only way to see other players' standings aside from the leaderboard, which is day, week and month, the only way to see that is in chat." (*Participant 1*)

A less direct link is in the use of chat to encourage player-led promotions. Promotions to the Scythe, Scout, Moderator and Mentor positions are now nominated and voted on by the player base and participant 1 stated the desire that players would pitch these roles to one another:

> "I would call sort of a community game mechanic, there are a lot more players who do achieve this ranked status and it becomes a kind of grass-roots levelling up system so rather than in most games there's a very clear path to get to the next level, in Eyewire it's like if you're active in the game and you're talking to people in the chat, then the players might say 'hey, are you interested in being a scout' and you're like 'I don't know what that is' and they explain it to you. It's almost like the players are pitching the younger players on this ranked class..." (*Participant 1*)

Although participant 1 did not draw an explicit link between this pitching and productivity, it should be noted that there are point and cube limits for scythes and scouts, such that any player pitched would have to meet the minimum requirements[3]. An additional indirect link was suggested between negativity in chat and player motivations. Participant 1 strongly felt that a positive chat atmosphere was important for the success of the project:

> "You know, you don't want a suspicious feeling, you want your community to be embracing and when someone new comes in and is kicking ass, you want your community to be congratulatory!" (*Participant 1*)

> "A citizen science project is competitive, it's not as competitive. It's like, collaborative and competitive at the same time. Like, we don't have trash-talking in Eyewire." (*Participant 1*)

Finally, participant 3 suggested that when assessing the overall efficiency of the project and individual players, consideration of chat plays an important role:

> "We do track how many cubes they play every kind of instance of gameplay. We have all their chat lines and stuff like that. So we have that as far as how busy players are." (*Participant 3*)

[3] `http://wiki.eyewire.org/index.php?title=Player_Roles`

6.7 Discussion

Overall, EyeWire is a social game-based activity and the comments and experiences shared by the team demonstrate a link between sociality – both interaction and communication – and the overall success of the project in terms of efficiency. Similar to the findings of Mugar et al. (2014), the team identify a link between sociality and facilitating VCS activities. Nevertheless, the idea of discuss as a way to facilitate complex tasks is relatively unique to the context of EyeWire. Neither Curtis (2015b) nor Jackson et al. (2014) describe such collaborative process, either between volunteers or with project scientists and with little collaboration is suggested by Crowston et al. (2018), in spite of the complex workflows associated with collaborative and co-created citizen science. As stated by Kim et al. (2014), EyeWire is a much more complex project than others within the VCS space, with significant variation in the difficulty associated with different assets. Furthermore, the automated algorithm used to spawn cubes is prone to spawning cubes if left to its own devices. If VCS projects are to move beyond the limited domains and task types described in chapter 4 and towards more collaborative models, greater interaction and co-operation will be required between project administrators and volunteers. While simple VCS tasks draw on distributed cognition as described by Smart (2017), through aggregated responses with participants unaware of each other, as tasks become more complex, there is a need for socially distributed cognition in which participants can exchange information with one another and share the cognitive load based on their own skills and knowledge.

In contrast with the opinions of project participants as described by Tinati et al. (2017a), the team also drew a clear link between discussion, motivation and engagement. On the one hand, none of the projects identified within this thesis require sociality in the strict sense. Just as suggested by participant 3, any VCS task is designed to be completed in isolation by anyone and then aggregated. This is reflected in the many users who successfully – albeit briefly – contribute to VCS projects without engaging in chat at all (Jackson et al., 2016b). Nevertheless, it is because of this use of redundancy and of repeatable microtasks that discussion becomes so crucial. Long periods of contribution can be repetitive, leading to participants becoming bored and leaving projects, yet it is long-term retention of participants that is most crucial to projects (Sauermann and Franzoni, 2015). Unlike projects like Galaxy Zoo, where the assets can be appealing in themselves, EyeWire participants face little variation between cells, except for in terms of difficulty (Raddick et al., 2010). According to the team, chat is explicitly used in EyeWire to prevent player burnout and to inject excitement into an otherwise potentially unappealing task, which aligns with how it is predominantly used by participants: in between, rather than during task contributions (Tinati et al., 2015a). This is particularly true for power players, who the team characterise as heavy and active chat users, further suggesting a link between productivity and sociality.

There is, however, a balance between sociality and the outputs of projects. Generally speaking, two broad classifications of success have been suggested in VCS – one offered by Cox et al. (2015b), focused on project outputs, publications and dissemination and a second offered by Wald et al. (2015), focused on ease-of-use, participant engagement and user experiences. These interview findings suggest that the EyeWire team have prioritised project outputs, particularly in terms of the quantity of submissions generated by the project, at times to the detriment of the user experience – for example the poor tutorial that has resulted in players leaving the game. This is not a phenomenon which is unique to EyeWire, as Wald et al. (2015) observed similar priorities across 20 different VCS sites. Even so, Wald et al. identify specific negative effects associated with such priorities, particularly in terms of participant learning and in the social elements of projects. This was reflected across the interview sessions, as while participants valued social factors such as sparking curiosity and allowing participants to experience science and learn together, they were nonetheless conscious of the need to prevent this to maximise contributions. This can most clearly be seen in the comments made by participant 1, that players "want to communicate with scientists," but that the team has had to prevent this as players become less motivated to do work due to the slow way in which science works. Productivity and sociality – along with the associated benefits such as learning – are therefore not always compatible in VCS and designers must carefully consider which of the two outcomes they value more.

A further interesting theme touched upon by the team was the importance of rewarding activity through extrinsic mechanisms such as points and badges. Initially this may not seem particularly surprising – EyeWire is, after all, a game that functions by rewarding points for contributions – but the interview process revealed that these extrinsic mechanisms appear to outweigh other significant factors. In spite of the importance of altruism and scientific interests as highlighted by chapter 4 and by Tinati et al. (2017a), even the most active and dedicated players were more driven by points, badges and leaderboard rankings than by scientific goals. In particular, the scythe complete feature demonstrates this clearly – it was not until players were more fairly compensated for their time that they began to engage with the activity. Returning to Downs' views as described in chapter 2, there is something of a contradiction here. According to Reisman's view of Downs' theories, Downs' five self-interest motivations are all "modes of superior-pleasing conduct" (Reisman, 1990). Yet in the case of EyeWire, if we take the view that the 'superiors' represent the project team, this is clearly not the case. Far from being motivated by a desire to please their superiors, top players appear to be motivated by the need to reinforce their own superiority, by maintaining their leaderboard positions and continuing to accumulate rewards.

Conversely, similar behaviours have been observed by Darch (2017) and by Woodcock et al. (2017), even in the absence of rewards. This suggests that, beyond extrinsic motivations, tensions exist between participants' perceptions of the scientific process and the realities that are experienced within projects. Ultimately the influence of participants' desires to contribute to science and the opportunity to interact with fellow participants and with scientists are of limited effectiveness. Participants will still favour more intrinsically or extrinsically motivating activities, even when scientists or project appeal to their altruism or their desire for interaction. On the one hand, this contradicts the findings of Lee et al. (2018) that participants shown messages stressing altruism and the opportunity to help science and scientists were more likely to visit and to contribute a project. However, the suggestion is not that participants are not motivated by a desire to help, but rather that what participants understand this assistance to involve and what scientists need are very different. The players in EyeWire who did not want to engage in Scythe complete, the players who were demotivated by the slow progress of research and the players angered by the need to classify simulated images described by Darch (2017) all had distinct views of science that did not align with the scientific process. Participant 1 suggests that sociality is of little benefit in this area, as in spite of the careful dialogue that was created with the scythes and the attempts to stress the importance of scythe completing, players ultimately only engaged when the task was changed to better appeal to their extrinsic motivations.

These findings are reminiscent of the prisoner's dilemma and zero contribution thesis as set out by Ostrom (1998) and Olson (1965). Previously, the complete function was available only to game masters and this led to delays and bottlenecks in the progress of the project, but also in the spawning of new cubes. This issue would be particularly problematic when the game masters were away from the project for any reason:

> "Actually let's put it this way. If we're not around, a lot less cells get done. That doesn't mean that less validations get done, because [players] are still submitting validations. But we're the only ones with the authority to mark cells as fully completed and add them into our tally for the month." (*Participant 6*)

As a result, scythe completion was added to attempt to reduce this bottleneck and increase the rate at which cubes are generated. This leads to something of a prisoner's dilemma for scythes, who must choose between maximising their own score by avoiding the scythe complete feature and in doing so, reducing the score of other players and the progress of the project as a whole or alternatively, completing cubes and increasing the scores of other players while potentially sacrificing their own score and leaderboard position. While EyeWire has collective elements, particularly in the form of competitions, the core gameplay is driven predominantly by maximising points and based on the team's observations, it is clear that top players are more driven by competition than the elements of collaboration. This was reflected in the statements made by interview participants, who noted that while they view the game as somewhat collaborative, they also view top players as more driven by points, rewards and status indicators such as leaderboard positions. Fundamentally, then, while EyeWire is much more social than the vast majority of projects presented in chapter 5, even the most active "social" players are driven predominantly by the opportunity to further their own individual achievements and in-game reputation.

On the other hand, this somewhat contradicts the findings suggested by Nov et al. (2014) that reputation features are perceived negatively by the community of Stardust@Home. Nov et al., however, draw a link between reputation and extrinsic motivations, suggesting that such motivations are not appealing for Stardust@Home volunteers. In contrast, EyeWire very much relies on extrinsic motivations, such as points and badges to drive long-term participation within the project. Just like Stardust@Home participants, EyeWire players do not report rewards as being particularly motivating, but do report finding the project fun and the core gameplay loop within EyeWire is very much dependent on gaining as many points as possible (Tinati et al., 2017a). While the project review findings suggest that more gamified projects are more social, this desire for points can also lead to more anti-social behaviours. Participant 1 drew a link between competition and aggression in some games – a behaviour which the team have tried to prevent in EyeWire. Moreover, the game masters had observed that players' desire to maximise their point score has driven them to avoid tasks which may be completed by less experienced players and new comers, due to the possibility of losing points. This suggests that task visibility and sharing within VCS can in fact have a negative effect on volunteer behaviour as participants are unwilling to work with players they deem to be inferior. This effect is clearly worsened by feedback and reward mechanisms, as shown by players reporting the concern that they will lose leaderboard rankings, points or reputation.

Arguably however, the greatest contradiction within the findings is the value suggested by the team of competitions and social events. When asked which factors most influenced their decision to participate in EyeWire, the community and competitions were mentioned by just 27 and 12 of 989 participants (Tinati et al., 2015a). Even so, participant 1 and to a lesser extent, participants 4, 5 and 6 described significant influences of periodic competition events on player numbers and retention and task completion, as well as potential effects on community engagement and task accuracy. These effects are particularly interesting given that EyeWire is by design somewhat competitive in itself. Since these competitions – particularly the marathon and task versus competitions – are the only opportunity for player to actively collaborate with one another (albeit weakly), these events also represent a form of sociality that is otherwise missing from VCS. They therefore represent an important angle for research and will be the focus of the quantitative analysis conducted within the next chapter.

6.8 Summary and Conclusions

In this chapter, I presented findings from three semi-structured interview sessions with six members of the EyeWire project team. Participants described the design process and history of EyeWire in detail, offering their opinions and anecdotal evidence for the inclusion of specific features, particularly rewards, social features and competitions within the project. The key findings of this chapter are:

- Productivity has been a central factor in influencing the design of EyeWire, particularly due to a need to achieve cost-effective but scientifically accurate results, while balancing the needs of players to ensure continued engagement.
- EyeWire predominantly draws on extrinsic motivations and the use of rewards – points and badges – to encourage continued participation and engagement from players.
- Sociality serves to break up repetitive tasks, to manage the game but also to facilitate tasks, particularly more complex tasks requiring collaborative action.
- Competition is central to the function of the game portion of EyeWire, with frequent competition events used to encourage player retention and drive productivity.

In the next chapter, building on these results I examine the effectiveness and impact of these competition events on productivity within EyeWire.

Chapter 7

Competitions – Sociality Within Task

Virtual Citizen Science is not truly a social phenomenon. As demonstrated in chapters 4 and 5 and echoed by the EyeWire team in chapter 6, Virtual Citizen Science relies on individual, randomly assigned crowdsourced tasks from which aggregated, more accurate responses can be inferred. Moreover, adding social activities – the opportunity to participate as a group or team – is a difficult or even impossible process. While FoldIt theoretically provides opportunities for 'group' and 'team' play, in truth these are no different from the usual gameplay activity, with the addition of a team chat service and the opportunity to improve on others players' completed solutions (Curtis, 2015b). Team play therefore offers no more collaboration than using the global chat service. Such forms of play further appear to be relatively rare – Curtis (2018b) suggests that many of the FoldIt teams are inactive or have retired from the game, with only around ten highly active teams, which are highly exclusive and which rarely accept new players. What projects *do* offer, however, is the opportunity to collaborate or compete with other participants while completing tasks. These activities were present in approximately a quarter of the sampled projects and in many cases offer an element of sociality to an otherwise individual-based task. Yet as seen in chapter 4 participants reported little motivation stemming from competition and the impact of competitions on participant engagement is a matter of some controversy within the literature.

In this chapter, I present results of statistical analysis and hypothesis tests to ascertain the association between social event competitions and player en-

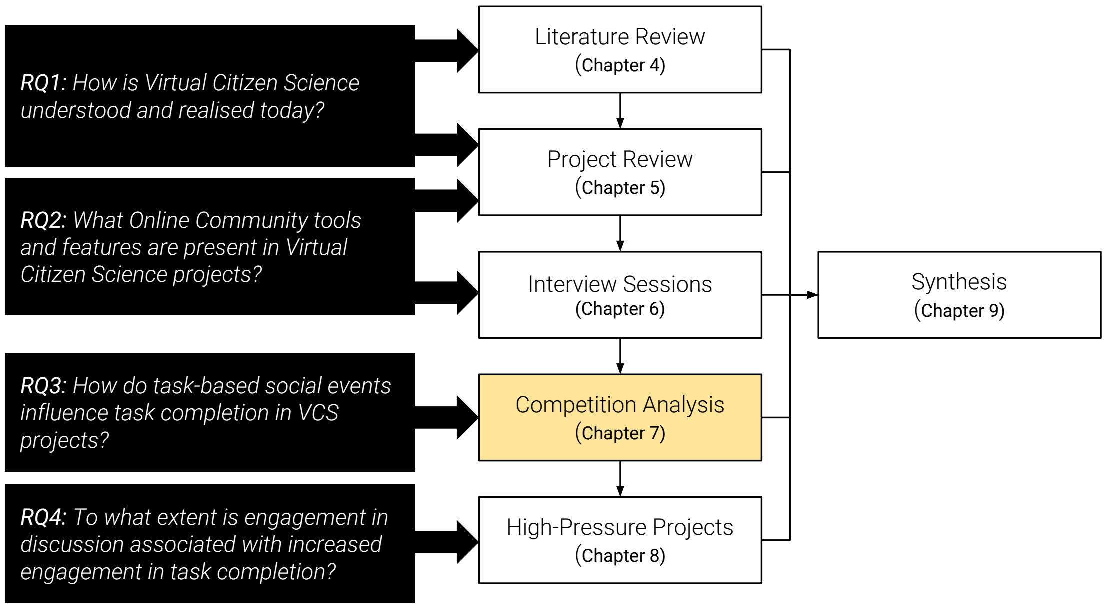


Figure 16: This study was partially motivated by – and draws on knowledge gained through – the interview process with the EyeWire team.

gagement within the EyeWire Game With a Purpose. Figure 16 illustrates how this stage succeeds from chapters 4, 5, and 6. After explaining the different types of competition and associated incentives, I present a set of hypotheses influenced by findings and conclusions from the literature review and interview process. Drawing on two datasets from the project, presented in chapter 3, I then test these hypotheses using the non-parametric Mann Whitney-U hypothesis test. These findings are then discussed in the context of the EyeWire project and in light of specific literature review and interview conclusions. This chapter ends with a conclusive summary regarding competitions in EyeWire.

7.1 Competition Types

EyeWire features two broad categories of competitions:

- *Minor* competitions occur regularly (fortnightly or bi-monthly), have thematic art-work and award players with points based on their own performance and the performance of the wider community or team to which players belong.

- *Major* competitions occur quarterly, with a week-long narrative arc, limited edition badges for participation and an immersive chat-focussed narrative experience. For example, during a ‘Whodunnit’ themed murder-

Table 11: Types of competition within the EyeWire project.

Type	Description	Availability
Accuracy Happy Hour	Players must complete cubes as accurately as possible, achieving an acccuracy of at least 80%	Major
Evil Cubes	Players must complete 12 very difficult cubes	Major
Happy Hour	Players earn bonuses based on points earned within a two hour period	Scheduled, Special
Marathon	Community must collaboratively complete cell in time limit	Minor, Major
Team Versus	Players must earn more points across 24 hours than the opposing team(s)	Minor, Major
Trivia	Players must answer questions through the chat interface before others	Major

mystery competition, game masters would drop hints to the identity of the killer (as well as red herrings) through the integrated IM-chat interface.

While the *marathon* and *team versus* competition classes occur both as minor and major competitions, the majority of major competition types can only occur quarterly. A summary of these competition types can be seen in Table 11. It should be noted that while the nature of competitions may change and while major competitions offer badges not available in minor competitions, the rewards for successful completion and participation do not otherwise vary between minor and major competitions. The bonus points awarded to a player are the same in both minor and major competitions. For the sake of this chapter, when comparing minor and major competitions, I will compare only the team versus and marathon competition varieties, as these have a duration of 24 hours rather than two hours, and to allow for a like-for-like comparison.

Although the *Team Versus* competition suggests team-based play, in reality, activities are no more team-based than in FoldIt. Players join a thematic 'team' by selecting their preference from two or more categories – examples include 'coffee vs tea' or during the launch of Pokémon Go 'Instinct vs Mystic vs Valor'[1]. However, players continue to complete the same activities before (cube tracing, scything and completing) on an individual basis, without opportunities for collaboration or team-based discussion. In fact, the team selected is largely irrelevant, with the exception of bonus point rewards which are offered to all

[1]Pokémon Go players must choose to join one of three in-game factions – Team Instinct, Mystic or Valor – at an early stage and the EyeWire teams paralleled this choice

members of the winning team. Nevertheless, the majority of awards are based on individual performance and do not take team performance into account:

- Players earn a 2,500 point bonus for earning 5,000 points as an individual.
- Players earn a 5,000 point bonus for earning 15,000 points as an individual.
- Players earn a 10,000 point bonus for earning 25,000 points as an individual.
- Players earn an additional 5,000 bonus points for every additional 25,000 points they earn as an individual.
- Each member of the winning team earns 10,000 bonus points.
- The highest scoring member of the winning team earns 5,000 bonus points.
- The overall highest scoring player (regardless of team) earns 10,000 bonus points.
- The second overall highest scoring player (regardless of team) earns 5,000 bonus points.
- The third overall highest scoring player (regardless of team) earns 2,500 bonus points.

During the *marathon* competition class, players race against time to complete an entire cell within 24 hours. This process entails completing thousands of cubes and as a result the goal itself is highly collaborative; no one player could possibly achieve the goal by his or herself. However, rewards for marathon participation are entirely based on individual performance. The one exception to this is that the opportunity to nominate and vote for an (internal) name for the cell is only offered to players if the goal is achieved within the 24 hour time limit. Nevertheless, the voting process itself is also restricted based on individual performance:

- Players who complete 20 cubes earn 2,000 bonus points.
- Players who complete 50 cubes earn 5,000 bonus points.
- Players who complete 150 cubes earn 20,000 bonus points.

- Players who complete 300 cubes earn 40,000 bonus points.
- Players earn 2,500 bonus points for every 150 cubes completed above 300 cubes.
- Players who complete at least 50 cubes earn the right to vote for an internal name for the completed cell.
- Players who complete at least 200 cubes earn the right to nominate an internal name for the completed cell.

The *Evil Cubes* competition class is part of the major competition cycle only, although it is only held semi-regularly. Players must complete 12 cubes which have been deemed to be extremely difficult (hence 'evil' cubes) by other members of the community. Rewards are offered based on overall accuracy relative to other players, although there is also a bonus for completing the 12 cubes. Players who complete the 12 cubes earn the right to nominate cubes for the next evil cubes competition. Two sets of cubes are run at any one time, so that players who nominated one set are unfamiliar with the other. Similar to Evil Cubes, *Accuracy Happy Hour* is an hour-long major competition activity, where players may complete as many cubes as they wish and must aim to achieve as high an accuracy score as possible. This score is based off of performance on a single cube and automatically selects the most accurate cube performed by each player who registers for the event; for example, a player who achieves 90%, 94%, 76% and 80% will receive an overall score of 94% for the hour. Unlike evil cubes, players are free to choose any cube during the event and performance is judged both on an individual-basis and relative to other players.

7.1.1 Hypotheses

As demonstrated in chapter 6, the EyeWire team ascribed a great deal of significance to the role that competitions play in EyeWire, viewing them as highly effective, particularly in terms of player productivity and project efficiency. This is particularly the case with major competitions, as demonstrated by the descriptions of Camp EyeWire and the EyeWire games. However, these views were predominantly a result of the *historic* outcomes of introducing social competition events. In particular, the majority of the team were unaware of having carried out specific statistical or numerical analysis of the outcomes of introducing specific competitions and competition features. Moreover, for the sake of hypothesis testing, it is not possible to conclusively state with certainty that competitions are the *cause* of any given outcome – merely that they are correlated with such an outcome (see section 7.1.2).

Although the hypotheses to be tested in this chapter are broadly influenced by comments made by the EyeWire team, it should be noted that these comments were not always specific and the team had not specifically analysed activity levels within EyeWire in drawing these assumptions. As a result and for the sake of this chapter, all hypotheses should be considered *exploratory*, being employed as a means to explore any association between holding competitions and outcomes in terms of engagement and activity within EyeWire. The intent with the formulation and testing of these hypotheses was not to assume that any of the two possibilities was more valid – that is, that the null hypothesis would hold true or not hold true for a given test. Rather, the formulation and testing of the null hypothesis was used as a method to explore the probability (in terms of p-value outcome) of an association between competition occurrence and specific outcomes.[2]

Hypothesis group A describes the correlation between the occurrence of competitions and the number of active players within EyeWire, as well as the number of cubes collected in total and per player. These hypotheses are derived from two distinct sources. The first is the view put forward by the EyeWire team that competitions are associated with an increase in overall site traffic and by extension, active player numbers. Since the team did not specifically distinguish between the impact of different competition types, this hypothesis group explores all competition varients – first all competition formats, then minor and major competitions and finally, a comparison of the major and minor competition classes.

In addition, results from the literature suggest a correlation between competitions and task completion rates. This association however, is nuanced, as described by Diner et al. (2018) and Laut et al. (2016), who suggest that the team versus competition class may be associated with a diminished correlation when considered in terms of per player contribution levels. Nevertheless, even the presence of extremely high performing peers was associated with greater than average participation levels for all players in these studies. Moreover, the

[2]Similar approaches have been used by Wallace et al. (2017), who used null hypotheses – however unlikely – as initial assumptions to test changes in publication rate. Similar work by Cheng et al. (2019) and Cheng et al. (2017), has gone further than Wallace et al., by avoiding commenting on the likelihood of an individual hypothesis altogether.

marathon class shares many of the features that make goal setting effective as described by Kraut et al. (2012) and Jackson et al. (2016a), particularly the presence of final and personal goals for participants to aim for. These findings lead to the following group A hypotheses:

H_A1 Competitions are associated with greater numbers of contributing players than on non-competition days.

$H_{A-null}1$ There is no significant association between the number of contributing players and the occurrence of competitions.

H_A2 Competitions are associated with a greater number of total cube submissions than on non-competition days.

$H_{A-null}2$ There is no significant association between the total number of cubes submitted and the occurrence of competitions.

H_A3 Competitions are associated with a greater number of cubes contributed per player than on non-competition days.

$H_{A-null}3$ There is no significant association between the number of cubes contributed per player and the occurrence of competitions.

H_A4 Minor competitions are associated with greater numbers of contributing players than on non-competition days.

$H_{A-null}4$ There is no significant association between the number of contributing players and the occurrence of minor competitions.

H_A5 Minor competitions are associated with a greater number of total cube submissions than on non-competition days.

$H_{A-null}5$ There is no significant association between the total number of cubes submitted and the occurrence of minor competitions.

H_A6 Minor competitions are associated with a greater number of cubes contributed per player than on non-competition days.

$H_{A-null}6$ There is no significant association between the number of cubes contributed per player and the occurrence of minor competitions.

H_A7 Major competitions are associated with greater numbers of contributing players than on non-competition days.

$H_{A-null}7$ There is no significant association between the number of contributing players and the occurrence of major competitions.

H_A8 Major competitions are associated with a greater number of total cube submissions than on non-competition days.

$H_{A-null}8$ There is no significant association between the total number of cubes submitted and the occurrence of major competitions.

H_A9 Major competitions are associated with a greater number of cubes contributed per player than on non-competition days.

$H_{A-null}9$ There is no significant association between the number of cubes contributed per player and the occurrence of major competitions.

H_A10 Major competitions are associated with greater numbers of contributing players than on days on which minor competitions occur.

$H_{A-null}10$ There is no significant association between the number of contributing players and the occurrence of major and minor competitions.

H_A11 Major competitions are associated with a greater number of total cube submissions than on days on which minor competitions occur.

$H_{A-null}11$ There is no significant association between the total number of cubes submitted and the occurrence of major and minor competitions.

H_A12 Major competitions are associated with a greater number of cubes contributed per player than on days on which minor competitions occur.

$H_{A-null}12$ There is no significant association between the number of cubes contributed per player and the occurrence of major and minor competitions.

Competitions in EyeWire are completely voluntary so a player who signs up for one competition will not necessarily sign up for another. It is therefore not logical to talk of the players who contribute to contributions, as any given player may be both a competitor and a non-competitor depending on the day or competition. It is nonetheless possible to consider those players who choose *never* to contribute to competitions, or at least, who did not contribute to competitions during the sample period – the *lurkers* characterised by participants 5 and 6 during the interview process. Notably, there is no passive benefit from competitions; those players who choose not to contribute to a competition, or who register but do no work, earn no rewards, badges or bonus points. It is likely on this basis that participants 5 and 6 described these users as lurkers and it is based on this characterisation that this group of hypotheses have been selected:

H_B1 Competitions are associated with greater numbers of active players from the non-competing group than on non-competition days

$H_{B-null}1$ There is no significant association between the number of active players from the non-competing group and the occurence of competitions.

H_B2 Competitions are associated with greater total numbers of cubes submitted from the non-competing group than on non-competition days

$H_{B-null}2$ There is no significant association between the total number of cubes submitted from the non-competing group and the occurence of competitions.

H_B3 Competitions are associated with greater numbers of cubes submitted per player from the non-competing group than on non-competition days

$H_{B-null}3$ There is no significant association between the number of cubes submitted per player from the non-competing group and the occurence of competitions.

H_B4 Minor competitions are associated with greater numbers of active players from the non-competing group than on non-competition days

$H_{B-null}4$ There is no significant association between the number of active players from the non-competing group and the occurence of minor competitions.

H_B5 Minor competitions are associated with greater total numbers of cubes submitted from the non-competing group than on non-competition days

$H_{B-null}5$ There is no significant association between the total number of cubes submitted from the non-competing group and the occurence of minor competitions.

H_B6 Minor competitions are associated with greater numbers of cubes submitted per player from the non-competing group than on non-competition days

$H_{B-null}6$ There is no significant association between the number of cubes submitted per player from the non-competing group and the occurence of minor competitions.

H_B7 Major competitions are associated with greater numbers of active players from the non-competing group than on non-competition days

$H_{B-null}7$ There is no significant association between the number of active players from the non-competing group and the occurence of major competitions.

H_B8 Major competitions are associated with greater total numbers of cubes submitted from the non-competing group than on non-competition days

$H_{B-null}8$ There is no significant association between the total number of cubes submitted from the non-competing group and the occurence of major competitions.

H_B9 Major competitions are associated with greater numbers of cubes submitted per player from the non-competing group than on non-competition days

$H_{B-null}9$ There is no significant association between the number of cubes submitted per player from the non-competing group and the occurence of major competitions.

H_B10 Major competitions are associated with greater numbers of contributing players from the non-competing group than on days on which minor competitions occur.

$H_{B-null}10$ There is no significant association between the number of contributing players from the non-competing group and the occurrence of major and minor competitions.

H_B11 Major competitions are associated with a greater number of total cube submissions from the non-competing group than on days on which minor competitions occur.

$H_{B-null}11$ There is no significant association between the total number of cubes submitted by the non-competing group and the occurrence of major and minor competitions.

H_B12 Major competitions are associated with a greater number of cubes contributed per player from the non-competing group than on days on which minor competitions occur.

$H_{B-null}12$ There is no significant association between the number of cubes contributed per player from the non-competing group and the occurrence of major and minor competitions.

An interesting consideration is whether there is a difference in association between the number of contributors and contributions from competing and non-competing players, on competition and non-competition days, particularly given the characterisation of non-competing players as lurkers. On the one hand, given that competitions are not a significant factor in terms of participant motivations in EyeWire according to Tinati et al. (2017a), it is possible that there is little or no difference in association between competitions and the total number of cubes submitted by the two groups, nor the number of cubes submitted by individual participants. Alternatively, assuming the view put forward by the EyeWire team is accurate and competitions are effective, then on competition days, it might be expected that there will be a greater association between player numbers and activity from the competing group. There is, then, something of a contradiction between the interview findings that competition participants are a small group of lurkers and the results demonstrated by Tinati et al. (2017a) that competitions motivate only a small number of players. The hypotheses for this group are as follows:

H_C1 Non-competition days are associated with greater numbers of active players from the non-competing group than from the competing group.

$H_{C-null}1$ There is no significant association between non-competition days and the number of active players from the non-competing and competing-group.

H_C2 Non-competitions days are associated with greater total numbers of cubes submitted from the non-competing group than from the competing group.

$H_{C-null}2$ There is no significant association between non-competition days and the total number of cubes submitted from the non-competing and competing group.

H_C3 Non-competition days are associated with greater numbers of cubes submitted per player from the non-competing group than from the competing group.

$H_{C-null}3$ There is no significant association between non-competition days and the number of cubes submitted per player from the non-competing and the competing group.

H_C4 Competition days are associated with greater numbers of active players from the non-competing group than from the competing group.

$H_{C-null}4$ There is no significant association between competition days and the number of active players from the non-competing and competing-group.

H_C5 Competition days are associated with greater total numbers of cubes submitted from the non-competing group than from the competing group.

$H_{C-null}5$ There is no significant association between competition days and the total number of cubes submitted from the non-competing and competing group.

H_C6 Competition days are associated with greater numbers of cubes submitted per player from the non-competing group than from the competing group.

$H_{C-null}6$ There is no significant association between competition days and the number of cubes submitted per player from the non-competing and the competing group.

H_C7 Minor competition days are associated with greater numbers of active players from the non-competing group than from the competing group.

$H_{C-null}7$ There is no significant association between minor competition days and the number of active players from the non-competing and competing-group.

H_C8 Minor competition days are associated with greater total numbers of cubes submitted from the non-competing group than from the competing group.

$H_{C-null}8$ There is no significant association between minor competition days and the total number of cubes submitted from the non-competing and competing group.

H_C9 Minor competition days are associated with greater numbers of cubes submitted per player from the non-competing group than from the competing group.

$H_{C-null}9$ There is no significant association between minor competition days and the number of cubes submitted per player from the non-competing and the competing group.

H_C10 Major competition days are associated with greater numbers of active players from the non-competing group than from the competing group.

$H_{C-null}10$ There is no significant association between major competition days and the number of active players from the non-competing and competing-group.

H_C11 Major competitions days are associated with greater total numbers of cubes submitted from the non-competing group than from the competing group.

$H_{C-null}11$ There is no significant association between major competition days and the total number of cubes submitted from the non-competing and competing group.

H_C12 Major competition days are associated with greater numbers of cubes submitted per player from the non-competing group than from the competing group.

$H_{C-null}12$ There is no significant association between major competition days and the number of cubes submitted per player from the non-competing and the competing group.

Due to the limited amount of data available regarding chat activity and the completion of other tasks, it is not possible to test the association between these activities and minor or major competition types. However, the EyeWire team did elude to difficulties in encouraging the completion of these secondary tasks during competitions, suggesting that this is a secondary aim of the competition process. It is therefore important to test there hypotheses and there is sufficient data to form hypotheses regarding the impact of competitions on these factors. Competitions are characterised by the EyeWire team as a community- and discussion-driven process, with a great deal of player excitement and the use of chat to deliver thematic details. With regard to additional activities that must be completed, the level of *scything* and *completing* required increases as the number of cubes submitted increases, although this relationship is more linear for completion than for scything[3]. These tasks are also only available to promoted, highly active players, rather than "lurkers". On this basis, we can devise hypotheses detailing an association between the amount of scything and completing performed and competitions:

H_D1 Competition days are associated with greater numbers of chat messages compared with non-competition days.

$H_{D-null}1$ There is no significant association between the number of chat messages sent and the occurence of competitions.

H_D2 Competition days are associated with greater numbers of chat participants compared with non-competition days.

$H_{D-null}2$ There is no significant association between the number of chat participants and the occurence of competitions.

[3]Each cube must be completed by three separate players. Scything can be completed by just one player and is only needed if a submission contains inaccuracies.

H_D3 Competition days are associated with greater levels of scything compared with non-competition days.

$H_{D-null}3$ There is no significant association between levels of scything and the occurence of competitions.

H_D4 Competition days are associated with greater levels of completion compared with non-competition days.

$H_{D-null}4$ There is no significant association between levels of completion and the occurence of competitions.

7.1.2 Correlation and Causation

The nature of the hypothesis test is such that it can be seen as synonymous, or at the very least, equivalent to a measure of correlation (Kraemer, 1975; Trafimow and Rice, 2009). Ultimately, the hypothesis testing process indicates the likelihood that a given observation would be viewed by chance and therefore the correlation between two given events as indicated by the null and non-null hypotheses to be tested (Trafimow and Rice, 2009). The hypothesis testing procedure alone is therefore insufficient to prove *causation* – that is, a given factor causes or leads to a given outcome, such as that the introduction of competitions results in an increase in task completion. Due to the nature of the data used within this chapter and the lack of specific A/B testing, it is not possible to indicate with certainty that the use of competitions is solely responsible for the observed associated effect. In essence, then, the outcomes of this chapter should be viewed only as indicating a potential *correlation* between the use of competitions in VCS and the observed results and not as an indication that competitions *cause* these observed outcomes.

7.2 Statistical Analysis

In this section, *non-competitors* refers to those players who chose not to engage with any competitions during the sample period. A set of descriptive statistics for each of the project types can be seen in table 12. Test statistics, conclusions and effect sizes for each of the hypotheses described in section 7.1.1 are described and summarised in table 14 for hypothesis group A, table 15 for hypothesis group B, table 16 for hypothesis group C and table 17 for hypothesis

Table 12: Daily mean statistics for each competition class (Dataset 1)

Type	Players	Percent Taking Part	Cubes	Chat Participants	Chat Messages	Percent from competitors
Accuracy HH	238.67	25.24%	5395.83	78.50	1606.00	70.21%
Happy Hour (Scheduled)	153.61	34.94%	4179.94	41.57	746.35	78.29%
Happy Hour (Extra)	133.15	38.44%	3971.08	N/A*	N/A*	N/A*
Marathon Min	146.75	35.05%	6300.75	37.00	766.00	87.86%
Marathon Maj	164.60	40.74%	5996.63	31.00	506.00	75.60%
Team Vs Min	152.59	36.58%	4429.09	45.36	955.20	77.75%
Team Vs Maj	191.20	29.05%	5887.38	43.80	911.40	81.37%
Non-Competition	139.18	N/A	3503.75	27.07	360.61	N/A

* Chat statistics unavailable for happy hour extra class due to a lack of instances of this class.

group D. Z-test statistics can be expressed as either a positive or negative value depending on the strength of the effect and which of the two conditions was associated with larger values. The Mann-Whitney U test requires that sample A be the larger of the two, so a negative Z statistic indicates that the tested values were larger in sample B than in sample A. While these conclusions are indicated within the four tables, for ease of interpretation of Z statistics, the non-competition sample size is always the largest, followed by the minor competition class, followed by the major competition class. In any test involving major competitions, a negative test statistic indicates that the values for the major condition were larger. Likewise, in any test involving the non-competition class, a positive test statistic indicates that the values for the non-competition condition were larger.

Table 14: Summary of hypothesis results – comparisons for total values. Z-crit for $p<0.01$ = 2.33 or -2.33

H	Condition	Z-Stat	*d*	Conclusion
$H_A 1$	Number of players (competition vs non-competition)	-5.1	0.55	Hypothesis confirmed. Competition days associated with higher total number of players than non-competition days ($p<0.01$, medium effect size)

(Continued)

H	Condition	Z-Stat	*d*	Conclusion
H_A 2	Total Cube Contribution (competition vs non-competition)	-6.99	0.79	Hypothesis confirmed. Competition days associated with higher total cube contributions than on non-competition days ($p<0.01$, medium/large effect size)
H_A 3	Cubes per player (competition vs non-competition)	-4.58	0.49	Hypothesis confirmed. Competition days associated with higher cube contributions per individual player than on non-competition days ($p<0.01$, small/medium effect size)
H_A 4	Number of players (minor vs non)	-5.66	0.72	Hypothesis confirmed. Minor competitions associated with higher number of players than non-competition days ($p<0.01$, medium effect size)
H_A 5	Total Cube Contribution (minor vs non)	-6.03	0.77	Hypothesis confirmed. Minor competitions associated with higher total cube contributions than non-competition days ($p<0.01$, medium effect size)
H_A 6	Cubes per player (minor vs non)	-3.37	0.41	Hypothesis confirmed. Minor competitions associated with higher cube contributions per individual player than non-competition days($p<0.01$, small effect size)
H_A 7	Number of players (major vs non)	-3.92	0.51	Hypothesis confirmed. Major competitions associated with higher total number of players than non-competition days ($p<0.01$, large effect size)
H_A 8	Total Cube Contribution (major vs non)	-5.78	0.79	Hypothesis confirmed. Major competitions associated with higher total cube contributions than non-competition days ($p<0.01$, medium/large effect size)
H_A 9	Cubes per player (major vs non)	-2.68	0.35	Hypothesis confirmed. Major days associated with higher cubes contributions per individual player than non-competition days ($p<0.01$, small effect size)
H_A 10	Number of players (minor vs major)	0.65	N/A	Hypothesis rejected. No statistically significant difference between player numbers on minor and major competition days ($p=0.52$)

(Continued)

H	Condition	Z-Stat	*d*	Conclusion
H_A 11	Total Cube Contribution (minor vs major)	-2.43	0.61	Hypothesis confirmed. Major competitions associated with higher total cube contributions than minor competition days (p=0.02, medium effect size)
H_A 10	Cubes per player (minor vs major)	-0.38	N/A	Hypothesis rejected. No statistically significant difference between the number of cubes contributed per individual player on minor and major competition days (p=0.70)

Table 15: Summary of hypothesis results – comparisons for non-competing players only. Z-crit for $p<0.01$ = 2.33 or -2.33

H	Condition	Z-Stat	*d*	Conclusion
H_B 1	Number of players (competition vs non-competition)	-5.1	0.55	Hypothesis confirmed. Non-competition days associated with higher number of players from non-competing group than non-competition days ($p<0.01$, medium effect size)
H_B 2	Number of cubes (competition vs non-competition)	-3.33	0.35	Hypothesis confirmed. Non-competition days associated with higher total cube contributions from non-competing group than non-competition days ($p<0.01$, small effect size)
H_B 3	Cubes per player (competition vs non-competition)	0.36	N/A	Hypothesis rejected. No statistically significant difference in the number of cubes contributed by individual non-competing players on competition and non-competition days (p=0.72)
H_B 4	Number of players (minor vs non)	-0.56	N/A	Hypothesis rejected. No statistically significant difference in the number of players from non-competing group on minor and non-competition days (p=0.58)
H_B 5	Number of cubes (minor vs non)	-0.07	N/A	Hypothesis rejected. No statistically significant difference in the total number of cubes contributed by non-competing group on minor and non-competition days (p=0.94)

(Continued)

H	Condition	Z-Stat	*d*	Conclusion
H_B 6	Cubes per player (minor vs non)	1.03	N/A	Hypothesis rejected. No statistically significant difference in the number of cubes contributed by individual non-competing players on minor and non-competition days (p=0.30)
H_B 7	Number of players (major vs non)	-3.44	0.45	Hypothesis confirmed. Major competitions associated with higher number of players from non-competing group than non-competition days ($p<0.01$, medium effect size)
H_B 8	Number of cubes (major vs non)	-3.16	0.41	Hypothesis confirmed. Major competitions associated with higher total cube contributions from non-competing group than on non-competition days ($p<0.01$, medium effect size)
H_B 9	Cubes per player (major vs non)	0.56	N/A	Hypothesis rejected. No statistically significant difference in the number of cubes contributed by individual non-competing players on major and non-competition days (p=0.58)
H_B 10	Number of players (minor vs major)	-3.65	0.92	Hypothesis confirmed. Major competitions associated with higher total number of players from non-competing group than on minor competition days ($p<0.01$, large effect size)
H_B 11	Number of cubes (minor vs major)	-3.36	0.83	Hypothesis confirmed. Major competitions associated with higher total number of cube contributions from non-competing group than on minor competition days. ($p<0.01$, large effect size)
H_B 12	Cubes per player (minor vs major)	-0.02	N/A	Hypothesis rejected. No statistically significant difference in the number of cubes contributed by individual non-competing players on major and minor competition days (p=0.98)

Table 16: Summary of hypothesis results – comparisons between competing and non-competing Players. Z-crit for $p<0.01$ = 2.33 or -2.33

H	Condition	Z-Stat	*d*	Conclusion
H_C 1	Number of Players (Competition)	-8.44	1.2	Hypothesis confirmed. Competition days associated with higher numbers of players from the non-competing group than from competing group on competition days ($p<0.01$, very large effect size)
H_C 2	Total Cube Contributions (competition)	13.3	2.76	Hypothesis confirmed. Competition days associated with higher total cubes contributions by players from competing group than from non-competing group ($p<0.01$, very large effect size)
H_C 3	Cubes per player (competition)	13.95	3.21	Hypothesis confirmed. Competition days associated with higher number of cubes contributed by individual players from competing group than from non-competing group ($p<0.01$, very large effect size)
H_C 4	Number of Players (Minor Competitions)	-4.07	0.88	Hypothesis confirmed. Minor competition days associated with higher total number of players from the non-competing group than from the competing group ($p<0.01$, large effect size)
H_C 5	Total Cube Contributions (Minor Competitions)	8.73	3.31	Hypothesis confirmed. Minor competition days associated with higher total number of cubes contributed by players from competing group than from non-competing group ($p<0.01$, very large effect size)
H_C 6	Cubes per player (Minor Competitions)	8.77	3.37	Hypothesis confirmed. Minor competition days associated with higher number of cubes contributed by individual players from competing group than from non-competing group ($p<0.01$, very large effect size)
H_C 7	Number of players (Major Competitions)	-4.22	2.0	Hypothesis confirmed. Major competition days associated with higher total number of players from the non-competing group higher than from the competing group ($p<0.01$, very large effect size)

(Continued)

H	Condition	Z-Stat	*d*	Conclusion
H_C 8	Total Cube Contributions (Major Competitions)	4.54	2.32	Hypothesis confirmed. Major competition days associated with higher total cube contributions from players from competing group than from non-competing group ($p<0.01$, very large effect size)
H_C 9	Cubes per player (Major Competitions)	4.95	2.92	Hypothesis confirmed. Major competition days associated with higher number of cube contributions by individual players from competing group than from non-competing group ($p<0.01$, very large effect size)
H_C 10	Number of Players (Non-competition)	-11.57	1.28	Hypothesis confirmed. Non-competition days associated with higher number of players from the non-competing group than the competing group ($p<0.01$, very large effect size)
H_C 11	Total Cube Contributions (Non-competition)	17.25	2.7	Hypothesis confirmed. Non-competition days associated with higher total cube contributions by players from competing group than from non-competing group ($p<0.01$, very large effect size)
H_C 12	Cubes per player (Non-competition)	18.34	3.3	Hypothesis confirmed. Non-competition days associated with higher number of cubes contributed by individual players from competing group than from non-competing group ($p<0.01$, very large effect size)

Table 17: Summary of hypothesis results – comparisons for total values. Z-crit for $p<0.01$ = 2.33 or -2.33

H	Condition	Z-Stat	*d*	Conclusion
H_D 1	Number of Chat Messages (competition vs non-competition)	-3.3	0.77	Hypothesis confirmed. Competitions associated with higher numbers of chat messages than non-competition days ($p<0.01$, medium effect size)
H_D 2	Number of Chat Participants (competition vs non-competition)	-2.64	0.60	Hypothesis confirmed. Competitions associated with greater number of chat participants than non-competition days ($p<0.01$, medium effect size)

(Continued)

H	Condition	Z-Stat	*d*	Conclusion
H_D 3	Number of Completions (competition vs non-competition)	-1.69	N/A	Hypothesis rejected. No statistically significant difference between total completions on competition and non-competition days (p=0.09)
H_D 4	Number of Scythe Actions (competition vs non-competition)	-0.21	N/A	Hypothesis rejected. No statistically significant difference between total number of scythe actions on competition and non-competition days (p=0.83)

In terms of the number of active players, the Accuracy Happy Hour class attracts the most active players, although on average only 25.24% of these choose to take part in the competition. It should be noted that this value is highly skewed because both Accuracy Happy Hour events take place on the same day in any given competitive period and this is also the day on which the competition launches and also always overlaps with either the initial Team Versus or Marathon event, depending on which of the two is held first. This can also be seen in the high number of cubes achieved on days on which the Accuracy Happy Hour takes place. Among the 24 hour activities, the marathon class is more effective than the team versus class in terms of the volume of cubes that are contributed, with on average 6300.75 cubes during the minor marathon and 5995.63 cubes during the major marathon. Again, though, this is skewed by the fact that on occasion minor marathons run two cells in a row, or add extra cubes after the first set are completed. The team versus class attracts a greater number of players, but a larger proportion of the active player base takes part in the marathon class. This is perhaps unsurprising given that the reward schedule for the marathon class offers greater rewards for individual players and achievement, particularly in the short term.

The number of chat messages sent by participants varies greatly, but on average is higher in all competition classes than during non-competition days. On any given competition day, on average, the vast majority of chat messages are sent by participants who have chosen to participate in that competition, with 87.86% of all chat messages sent during minor marathon competitions produced by players who have chosen to contribute to the marathon. The number of chat participants is also higher during all competition classes, although this value is very low outside of competitions, with just 27.07 participants, rising to just

31.00 during the major marathon competition. Days on which Accuracy Happy Hours occur result in the largest number of chat participants, with 78.50 players contributing to chat, but this is also likely because these days are the days on which the major competition events occur, rather than a direct result of this competition class.

7.2.1 Participation

In line with the motivations reported by Tinati et al. (2017a), only a small minority of EyeWire players choose to take part in competitions. During the sampled period, a total of 10,296 players contributed at least one cube to EyeWire. EyeWire does not allow for anonymous contributions and so this reflects all users except those who did not complete the tutorial process, or who otherwise logged into the project but did not perform any work. During this period, 143 competitions took place and 494 players earned at least one point or contributed at least one cube to at least one of these competitions – just 4.8% of the total number of players active during this time. The mean number of competitions entered by the 494 players is 6.50, but this value is skewed by a small number of highly active players and over half of players (267 users) who engaged in just one competition.

7.2.2 Player Numbers

Competitions are associated with a significant increase in player numbers within EyeWire. The number of players who contribute to EyeWire is significantly higher on days on which competitions are held than on days on which they are not ($p<0.01$, $d=0.55$). This is true of both minor and major competitions ($p<0.01$, $d=0.88$; 0.72). There is, however, no statistically significant difference between the number of players on minor and major competition days ($p=0.52$). Even among those players who chose not to contribute to competitions during the sample period, the number of active players is higher on competition days than on non-competition days ($p<0.01$, $d=0.55$). Despite this, when comparing minor competition days with non-competition days, test results do not allow for the rejection of the null hypothesis ($p=0.58$). Given that major competitions attract significantly more players from the non-competing group than minor competitions and the effect size for this test is large ($p<0.01$, $d=0.92$), it can be concluded that there is no statistically significant difference in non-competing player numbers between non-competition and minor competition days. Competitors, however, represent a minority of the active EyeWire player base across all days, reflecting the low proportion of players who actively contribute to competitions. This effect is weaker during competitions than outside of competition periods (d=1.2; 1.28). Even so, the effect is significantly stronger during major competitions, reflecting the increase in player numbers among the non-competitors group (d=2.0).

7.2.3 Cube Contributions and Productivity

Competitions are associated with a significant increase in the number of cubes contributed by players, both as a whole and individually, although this effect is larger when considered across the community than when examining contribution rates per player ($p<0.01$, d=0.79; 0.49). Both major and minor competitions are associated with higher numbers of total cube contributions, although the effect size is slightly higher in major competitions than in minor competitions (d=0.79; 0.77). This is reversed in terms of the number of cubes contributed per player, where the effect sizes are smaller, with the effect size in minor competitions greater than major competitions (d=0.41; 0.35). There is no statistically significant difference in the number of cubes contributed per player between minor and major competitions, but significantly more cubes are contributed in total in major competitions (d=0.61). However, the p-value for this condition is slightly higher than in other conditions, reflecting a 98% rather than greater than 99% chance that this result would not be seen by chance.

Among the non-competing group, the total number of cubes submitted is greater during competitions than non-competition periods (d=0.35). When tested separately, major competitions result in significantly more cube classifications than non-competition periods, but minor competition periods do not. Major competitions also attract significantly more total cube contributions than minor competitions and so it can be concluded that minor competitions have no significant effect on the number of cubes contributed by non-competitors. There is no statistically significant difference in the number of cubes completed per player among non-competitors across any of the four conditions.

Competitors contribute significantly more cubes than non-competitors on competition days, both minor and major and on non-competition days as well. Although this effect size is extremely large in all cases, it is greatest in the case of major marathons (d=3.31) and weakest among major competitions (d=2.32). When considering the number of cubes contributed per player, once again competitors achieve significantly higher productivity than non-competitors, with this effect greatest on minor competition days (d=3.37) and weakest on major competition days (d=2.92).

7.2.4 Chat and Additional Tasks

Competitions are associated with a significant increase in the number of chat messages sent by players and the number of chat participants over non-competition days, both of medium effect size (d=0.77; 0.6). In contrast, no statistically significant difference is observed with regard to scything and completion between competition and non-competition days. No consideration was given to the impact on the number of scything and completing participants due to the low number of promoted players and the fact that more so than with cubes, the vast majority of scything and completing are completed by a small number of very active players. This echoes the issues raised by the EyeWire team in chapter 6 and suggests that completion remains a relatively unengaging activity, even after changes to the reward schedule and emphasis of the importance of completion and its altruistic value for the project itself and for neuro-scientific research. Moreover, since three completions are required for every submitted cube, the number of completions required increases linearly with every additional cube.

7.3 Discussion

In spite of the low number of participants who described competition as a motivating factor for their participation in EyeWire according to Tinati et al. (2017a), competitions are associated with significant increases in the number of contributing players, the productivity of players and the overall number of cubes contributed. There are, however, additional motivational factors involved in competitions, particularly extrinsic motivations, as competitions are a simple way to earn greater rewards without requiring a significant change in activity. Even so, while interview participants characterised the community as predominately driven by points, it bears consideration that just a small minority of participants have contributed to competitions and in the majority of cases, these participants have engaged in just one event. If participants truly are so driven

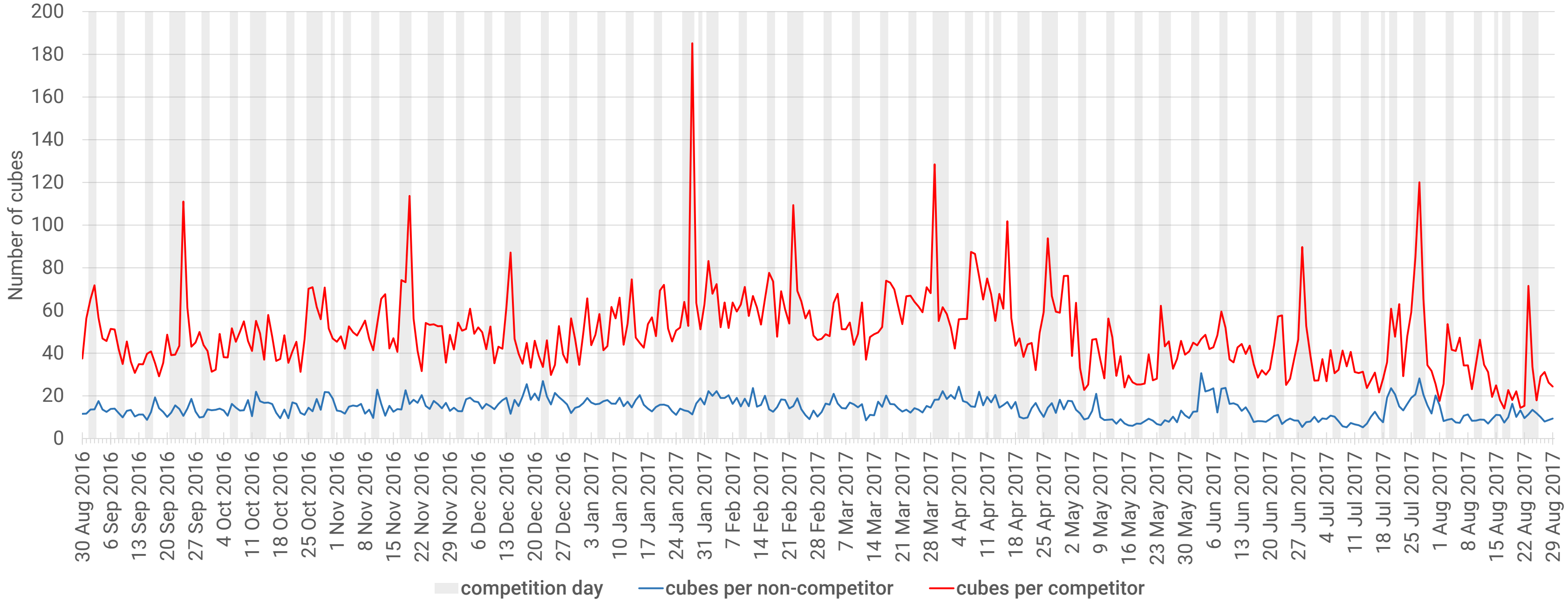


Figure 17: Cubes per non-competing and competing players on competition and non-competition days. Grey highlighted periods correspond to competitions.

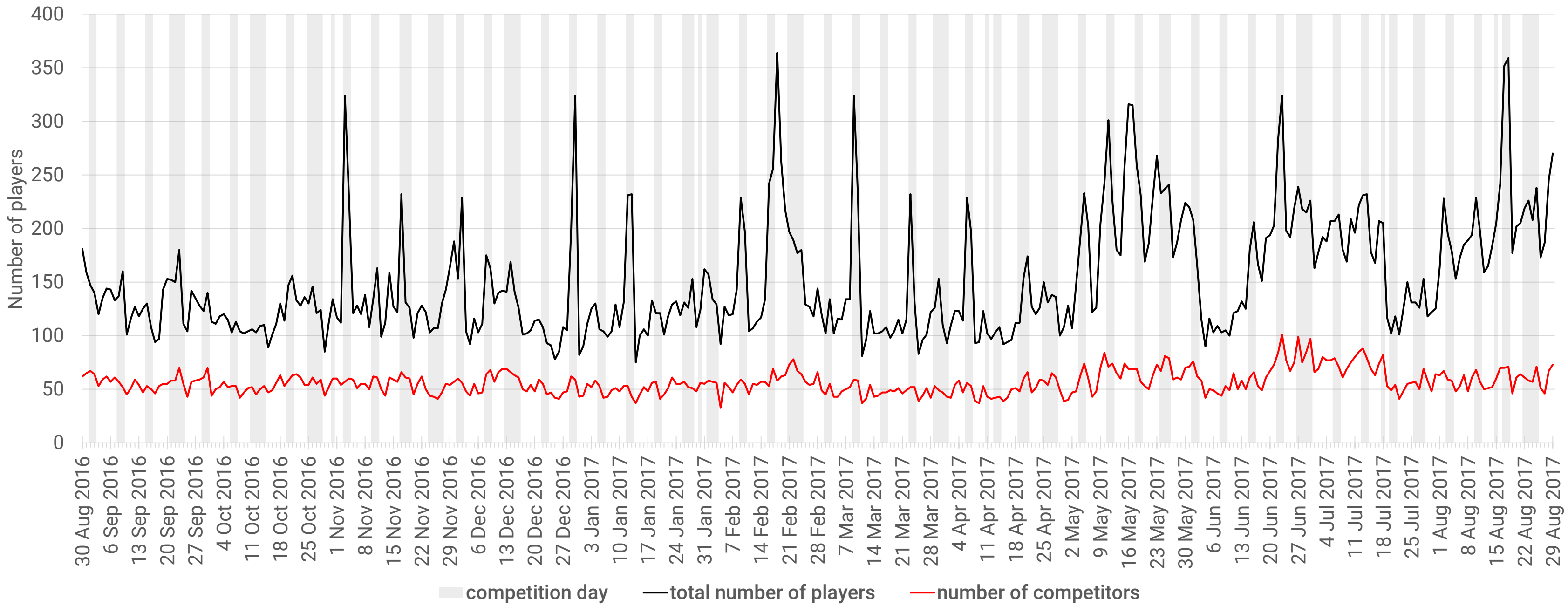


Figure 18: Number of non-competitors and total number of players on competition and non-competition days. Grey highlighted periods correspond to competitions.

by points, it seems counter-intuitive that they would pass up an opportunity to earn relatively simple bonuses while they play. There are, after all, no barriers to entry and no sacrifices required for competition participation. If a participant performs poorly, he or she cannot lose points and while some forms of competition restrict the activity a participant can engage in (for example, no scything or completions count towards marathons), this will be irrelevant for the majority of participants who will not have received the necessary promotion to complete these tasks.

The increases associated with competitions stretches beyond those who choose to participate within competitions. As seen in table 15, the number of non-competing players is greater during competition periods than non-competition periods and this effect is largest during major competitions. Similarly, cube contributions from this non-competing group are also larger during competitions, particularly major competition periods. The reason for this is unclear. At first glance, it may appear odd that non-competitors would choose to sign in during major competitions in significant numbers. Players who choose not to contribute in a given competition do not earn bonuses, rewards or badges and there is little to spectate, with the exception of the leaderboard scores, which are present during both minor and major competitions. For non-competing players the only benefit to participating in the project during a major competitive period is the opportunity to view and engage with the thematic elements associated with the competition, the majority of which occur through the project blog rather than the chat system. In the case of the EyeWire Games competition, there are no narrative features for players to experience and yet according to the EyeWire team, these have been highly successful. There is however, a strong element of communal spirit that occurs in the chat system during competitions; as participant one stated "chat is blowing up. Everyone's excited." The opportunity, then, to experience a communal event with other players can be seen as an exciting opportunity even for those players who are less active and who do not choose to actively participate within the event.

Nevertheless, this does not appear to be the case for competing players. Major competitions do not attract significantly more players, nor result in more cubes contributed per player than minor competitions. Notably, however, when considering *all* major competition days, including those on which the Accuracy Happy Hour and Evil Cubes events are held, major competitions do attract significantly more players than minor competitions and the effect size of this impact is large (Z=-3.5, p<0.01, d=0.88). This reflects the popularity of the opening and closing periods of competitions, during which narrative elements are launched and activities such as player awards and promotions occur. The popularity of the opening day suggests that the communal experience is partially appealing to competing players, but overall, these findings demonstrate that they are predominantly driven by extrinsic factors and peer-recognition. Rewards are no greater for major competitions than for minor competitions and as a result, competing players show no preference for either competition variety.

There are, potentially, other factors that have influenced these findings. Specifically, major competitions are accompanied by email campaigns, but the timing of these campaigns is inconsistent, occurring on the day of a competition launching, the day or days leading up to the event or even, in some cases, a full week before. Similarly, the nature of these emails also changes, with some describing only the upcoming event and others giving a full run down of EyeWire news, as can be seen in figure 19. As demonstrated by figure 18, the number of players active on any given day tends to be peak at around the time of competitions and in all cases, peaks in player numbers are brief. On this basis, it is unlikely that the email campaigns attached to projects have a sufficiently significant effect on player numbers to impact the findings outlined here.

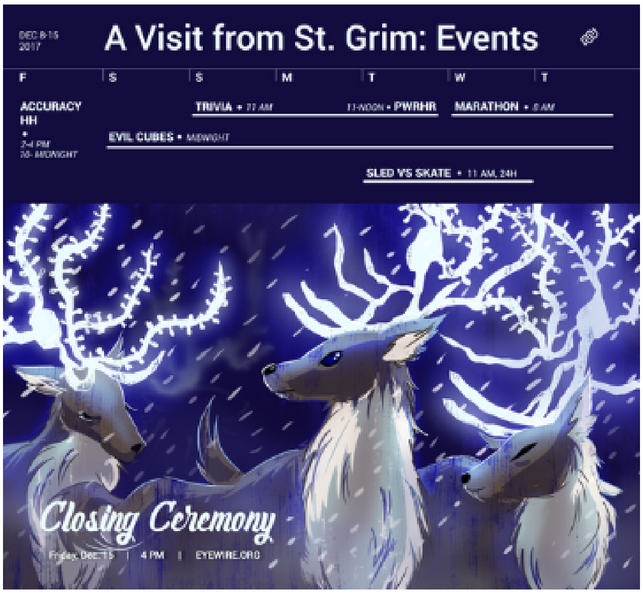
While competitions may have a beneficial effect in terms of the volume of cubes contributed and the number of players, they also pose significant issues in terms of project efficiency. Arguably more important than simply contributing cubes is the community-led quality assurance process of scything and completing. Ensuring these tasks are completed is already a significant issue within EyeWire. As the game masters described during the interview, during periods where the administrators are busy or on holiday, a backlog builds up due to a lack of cube completions. Competitions are not associated with a significant variation in the number of completions or scythings that occur and this also suggests a backlog: as the number of cubes that are completed increases, so more work is created to scythe and to complete. While it is true that not all players can engage in this task, this is not the cause of the issue as scything and completing was performed almost exclusively by players from the competitor group. The team versus competition also recognises scything and completing as contributions and these raise participant and team scores. As interview participants pointed out, however, these tasks remain less efficient in terms of rewards and for those highly motivated by competition, cube completion is the more efficient

Eyewire, a game to map the brain, is here with some holiday cheer for you!

A Visit from St. Grim

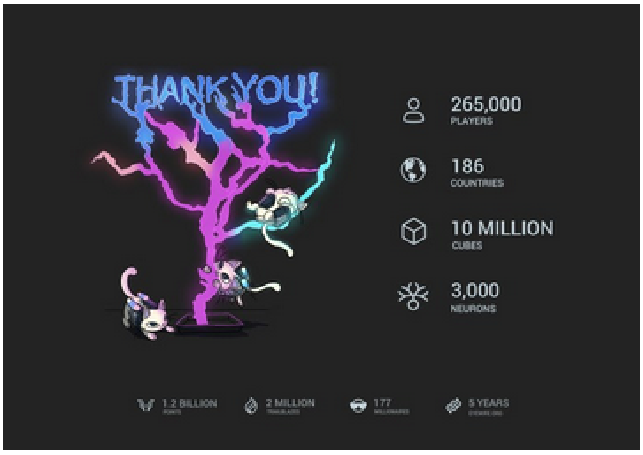
Twas the night before Grim-mas and all through the house,
Eyewirers were tracing the brain of a mouse.
Now's the time for excitement, let us start our next mission,
And get all bundled up for a festive competition!

The fun begins this Friday 12/8.



Eyewire's 5th birthday!

Eyewire launched on December 10th 2012, which means that we've been going strong for 5 years now. Wow! Thank you for being a part of the Eyewire community. Check out this post to see a ton of stats plus a timeline of Eyewire's major accomplishments.



Eyewire is hiring!

Are you a full stack developer with a passion for games and an interest in the brain? Join our Boston team to help bring the future of neuroscience, citizen science, and gaming to the next level! Send your resume, code sample, and cover letter to jobs@eyewire.org. Details about the position here.

Figure 19: Example of EyeWire Competition campaign email.

way to earn points.

This conflicts somewhat with earlier findings. Within the literature, autonomy has been highlighted as a key motivational affordance across a variety of human computation initiatives (Lukyanenko et al., 2016; Prestopnik and Tang, 2015). Conversely, autonomy in competitions is associated with a trade-off in terms of engagement within EyeWire. The more autonomous team versus competition class – during which participants can earn points from any cube and, where applicable, from bonuses, tutorials and promotion-dependent activities – attracts more players and competitors, but on average resulted in lower numbers of cube contributions when compared with the marathon class. Further findings from the literature suggest that autonomy can be associated with chat and the opportunity to self-organise or to exercise different skills. This view was also put forward by the EyeWire team, in stating that *"chat makes up for the dryness of the task."* Yet the mean number of chat messages sent during the autonomous marathon class is significantly lower than during the team versus task.

Conversely, Kraut et al. (2012) and the EyeWire team have both observed that participants in online communities respond better to specific, concrete goals and the team note this is particularly the case in VCS, where overarching scientific goals are too high-level and progress too slowly. As participant 5 pointed out, curing cancer does not function as an achievable target. Even so, the use of specific, time-coupled goals as suggested by Kraut et al. (2012) has unintentional consequences in VCS, as participants become more focused on the task and less focused on social interaction – reflecting the finding of Tinati et al. (2015a) that highly active players use the silence command while completing cubes. While this is clearly beneficial for productivity among these players, given the importance of social experiences for other, less competitive players, careful consideration is required for how best to balance the needs of different elements of the EyeWire community. During the interview phase, participants placed a great deal of importance on ensuring a positive atmosphere for EyeWire players. As participant 1 described during the interview process, it is important that players – particularly new players – feel accepted and encourage one another to achieve more, similar to the experience of participants in FoldIt and Planet Hunters (Curtis, 2015b; Jackson et al., 2014).

7.4 Summary and Conclusions

This chapter has presented the results of an analysis of player numbers and contributions within the Virtual Citizen Sience Game With A Purpose, EyeWire. These competitions add a degree of sociality to projects by allowing participants to contribute in loosely connected teams, or as a collective towards a common goal. To summarise, the main findings of this chapter are:

- The implementation of competition events is associated with increased active player numbers, total task submissions and task submissions per individual.
- This effect extends beyond those who choose to take part, with non-competitor numbers increasing and greater task submissions from this group as a whole on competition days.
- Competitions are similarly associated with increased discussion participation from both groups, with an increase in chat comments sent and total chat participants.
- However, no association is seen between competitions and additional tasks which are not rewarded through the competition schedule.
- Similarly, all effects are restricted only to competition days and diminish rapidly after the competition is complete.

In the next chapter, I will explore the extent of the relationship between discussion and task contributions in VCS by exploring contributions to two high-pressure projects.

Chapter 8

Sociality and Productivity – High Pressure Projects

Although various sources – including the EyeWire team, Tinati et al. (2014) and Tinati et al. (2015a) – have suggested that the most active talk participants are the most active task participants, the extent and validity of this relationship has yet to be tested. In particular, there has been little consideration of how periods of high-productivity influence – or are influenced by – discussion activity within projects. In this chapter, I draw on two full VCS projects completed over the course of 48 hours with specific deadlines, which represent high-pressure situations for VCS participants, to understand how discussion ocurs within these projects. I begin by outlining the projects and the context in which they arise, as it represents something of a departure from traditional VCS projects which occur over the course of months or years. This is followed by descriptive statistics regarding average contribution levels within each of the two projects and an analysis of how contributions vary in each project other time. To understand and quantify the relationship between talk and task, a correlation of these factors in each project is then carried out, followed by hypothesis testing and calculation of effect sizes to expand on and clarify these findings. The chapter ends with a discussion of the results described here and a comparison with previous chapters' findings, as well as a summary and conclusion of the main points.

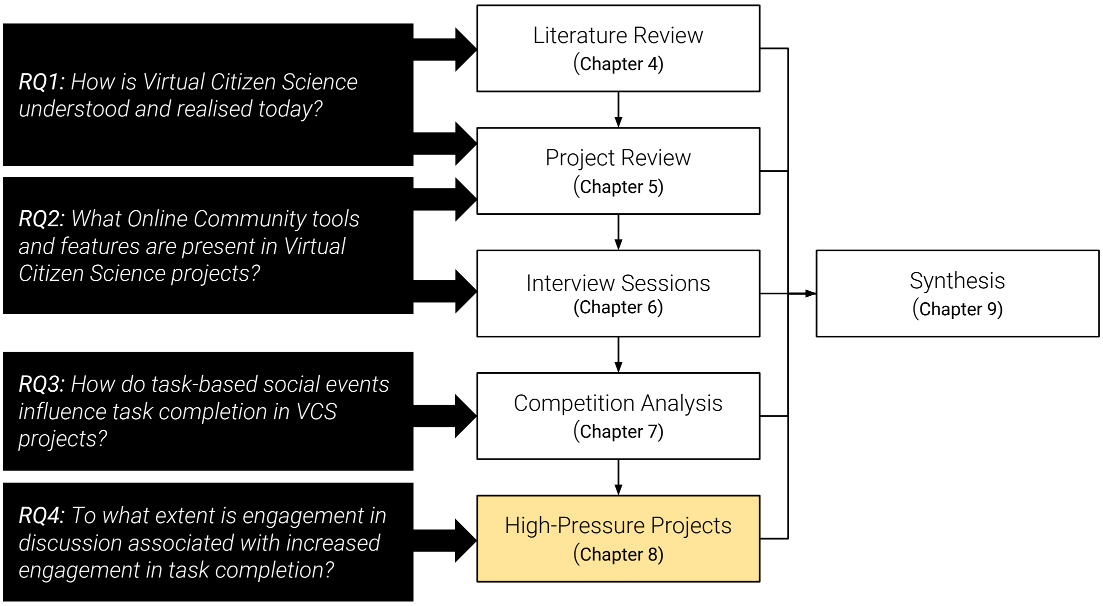


Figure 20: This study uses quantitative methods that build on findings from the literature review and project review.

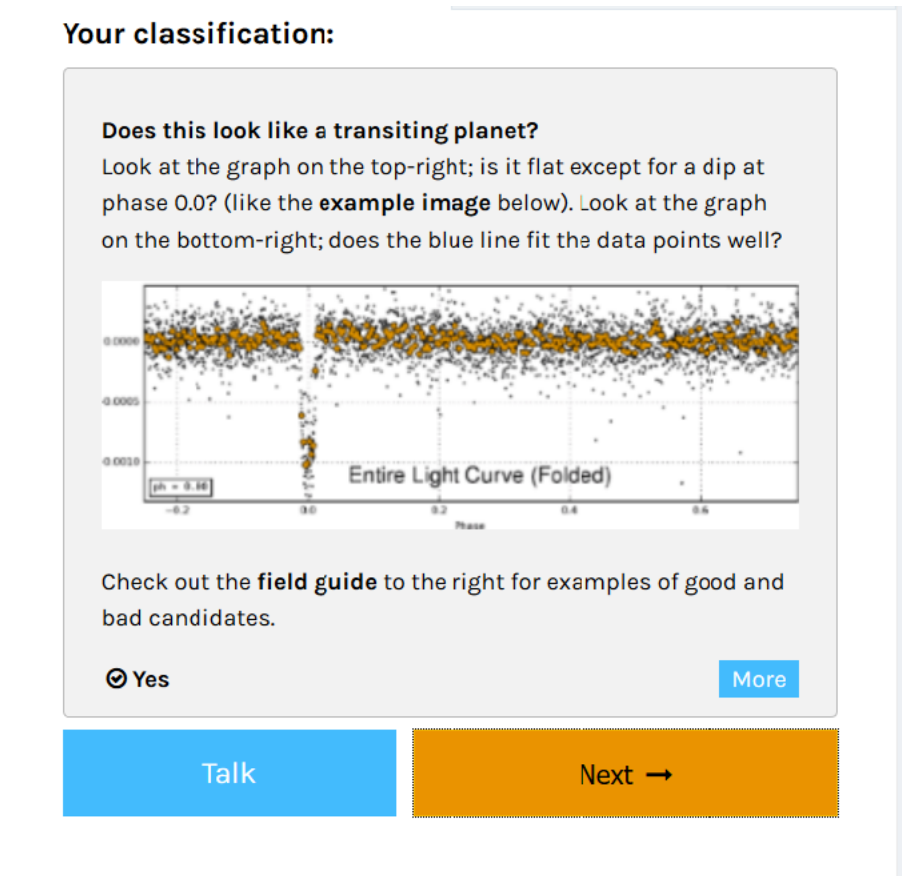


Figure 21: Classification submission screen for the Exoplanet Explorers project, showing the Talk prompt and task submission button.

8.1 Project Contexts

Each year, in partnership with BBC Stargazing Live, the Zooniverse launches a 48 hour campaign held across three days, to generate as many classifications from participants as possible, as described in chapter 5 (Aye et al., 2018). After 48 hours, scientific results generated directly through the campaign are announced live on air (Aye et al., 2018). Initially, as described, these broadcasts have showcased existing Zooniverse projects such as Planet Four, Planet Hunters and Spacewarps (Masters et al., 2016). The partnership with Stargazing Live and use of live broadcasts has proven highly successful, driving far more participants to contribute to these projects than would otherwise be expected (Curtis, 2018a; Masters et al., 2016). In recent years, however, rather than using existing projects, new projects have been launched using Stargazing Live to drive initial participation, such as *Snapshot Supernova* (Trouille, 2016). This chapter looks at two such projects – Exoplanet Explorers and Planet Nine – launched across ABC and BBC Stargazing Live in 2017, with the intention that participants would complete all necessary classifications in just 48 hours.

Both Exoplanet Explorers and Planet Nine are astrophysics projects, aiming to discover evidence for new planets. While both technically correspond to the *categorising* task type as described by Dunn and Hedges (2013), each uses a far simpler task workflow than the projects described in chapters 4 and 5. Unlike Snapshot Serengeti which features in excess of 90 categories, or Galaxy Zoo which features a branching decision tree (Prestopnik and Crowston, 2012; Swanson et al., 2015), participants must simply make a binary yes/no choice based on the appearance of the asset presented to them, as shown in figures 22 and 23. Planet Nine draws on image assets, while Exoplanet Explorers uses graphs, with – as demonstrated by figure 21 – comparative examples of target assets to simplify the task expected of participants. Planet Nine ran from the 31st of March until the 2nd of April and was ultimately unsuccessful in finding evidence for the existence of a hypothesised ninth planet in our solar system, but participants did rediscover a number of existing known astrophysical phenomenon using the Talk system [1]. Exoplanet Explorers ran the same week, from the 4th to the 6th of April and was more successful. In total volunteers found 184 potential exoplanetary candidates, either through classifications[2] or

[1]https://www.zooniverse.org/projects/skymap/planet-9/talk/793/243564?page=1
[2]https://www.zooniverse.org/projects/ianc2/exoplanet-explorers/about/results

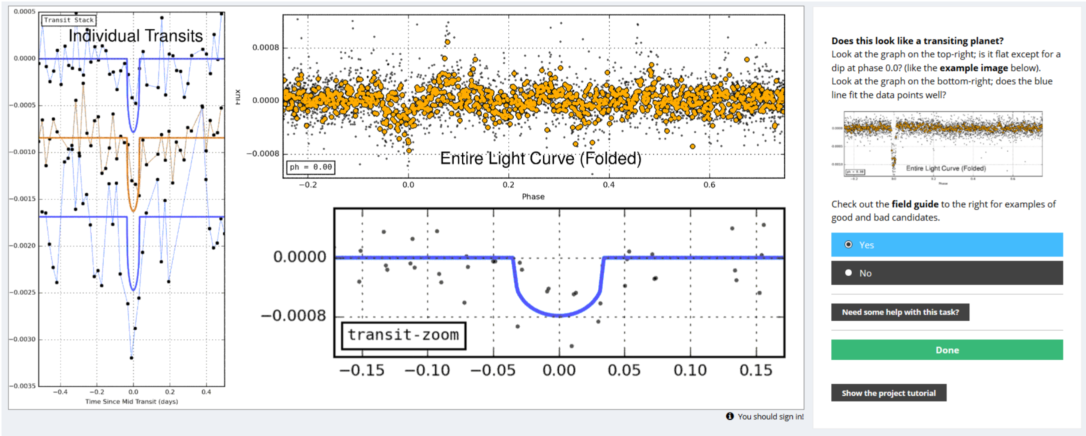


Figure 22: Exoplanet Explorers interface, showing image asset and yes/no asset classification task

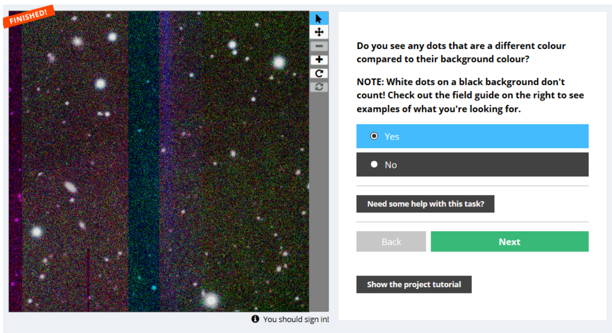


Figure 23: Planet Nine interface, showing image asset and yes/no asset classification task

by sparking and contributing to discussions through the integrated chat platform 'talk'[3], in a manner similar to the discovery of the Hanny's Voorwerp or the Galaxy Zoo Green Peas (Cardamone et al., 2009; Lintott et al., 2009).

8.2 Descriptive Statistics

Descriptive statistics for the two projects can be seen in table 18. In spite of the difference in the total number of tasks completed, the mean number of hourly

[3] https://blog.zooniverse.org/2017/04/07/stargazing-live-2017-thank-you-all/

Table 18: Descriptive statistics for Exoplanet Explorers and Planet Nine, including mean hourly statistics, total contributions and volunteer numbers.

Name	Mean Task	Mean Talk	Total Task	Total Talk	Volunteers	Chat Participants
Exoplanet Explorers	29376.86	189.83	2762043	17,844	8,164	1,155
Planet Nine	32647.56	193.02	4570659	23,355	20,346	2,526

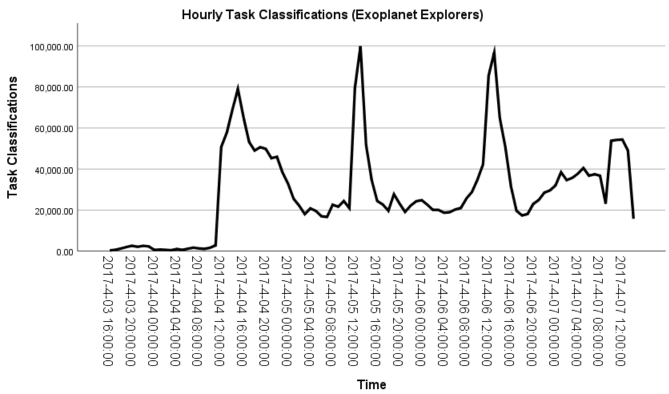


Figure 24: Line graph showing hourly contribution counts within the Exoplanet Explorers Project

task and talk contributions for each project is relatively similar. This is partially skewed, as Exoplanet Explorers was soft launched over 24 hours prior to the first Stargazing Live broadcast, while Planet Nine was not retired immediately and remained available (with a warning that classifications had been completed) after its ‘retirement’ – see *contributions over time*. Exoplanet Explorers attracted less than half as many registered users as Planet Nine, although a larger proportion of Exoplanet Explorers engaged with Talk in Exoplanet Explorers. This reduced engagement may be related to the fact that this was the first time that ABC Stargazing Live was shown and thus, the first time that a second Zooniverse project was launched in this period. A significant number of classifications were completed by unregistered users in each project – 658,854 in the case of Planet Nine and 217,595 in the case of Exoplanet Explorers. No data were available to allow these classifications to be associated with individual users, or to allow the grouping of classifications based on, for example, similarities. For this reason, these were removed from the sample prior to analysis.

8.3 Contributions Over Time

Exoplanet Explorers was given a soft launch at approximately 16:00 on the 3rd of April 2017, but was not featured on ABC Stargazing Live until 10:00 (20:00 in the Australian GMT+10 timezone) on the 4th of April 2017. Although initial task completion rates were extremely low, these rates increased significantly in the run up to the live broadcast. An initial peak of 79,086 classifications was made in the hours following the first broadcast and the ‘official’ launch of the project. Following this initial broadcast, the classification rate diminished somewhat rapidly over the course of 10 hours, reaching a low of 16,670 classifications.

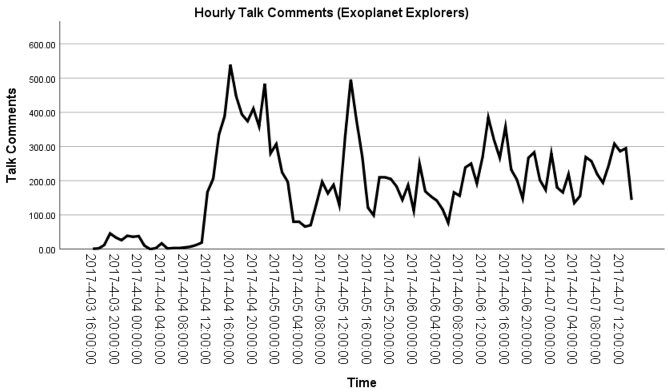


Figure 25: Line graph showing hourly talk comment counts within the Exoplanet Explorers Project

With the second broadcast event, classifications increased significantly, reaching an overall peak of 99,868 classifications in the hour following the broadcast, before once again diminishing to 18,689 classifications at a far more rapid pace than following the initial broadcast event. After 24 hours, the classification rate once more rose after the final broadcast event, initially deteriorating again before slowly and steadily rising over the final 24 hours of the project lifespan.

Talk comment contributions similarly varied greatly, with peaks corresponding to the completion of broadcast events. However, unlike task contributions, the greatest peak in talk comments occurred after the first broadcast event, with a second smaller peak after the second event and a final, significantly smaller peak after the third and final broadcast event. Notably, however, talk comments show something of a positive trend, with a gradual increase in the number of hourly talk comments made with the exception of the final hour of the projects' lifespan. Talk comment rates were also far more prone to variations than task contribution rates, with frequent small but significant increases and reduction in hourly talk comment totals. Similarly, during the soft launch period, Exoplanet Explorers attracted disproportionately more talk comments than task contributions.

In contrast with Exoplanet Explorers, Planet Nine launched on the Zooniverse platform just prior to the initial BBC Stargazing Live broadcast at 20:00 GMT and generated immediate classifications from participants, with the greatest classification rates in the hours following the initial broadcast during which time participants contributed 301,680 classifications in an hour - over three times the peak contribution rate of Exoplanet Explorers. In turn, after approximately 3 hours, classification rates significantly dropped off at a more

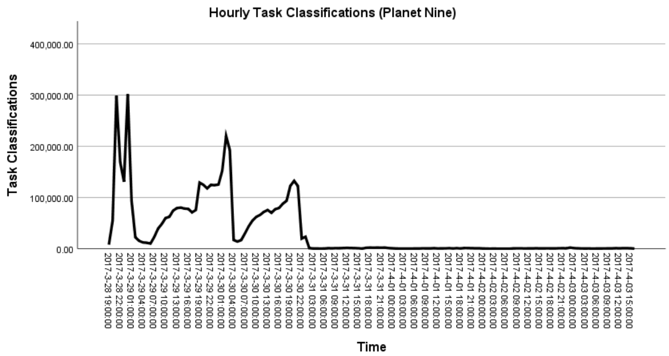


Figure 26: Line graph showing hourly contribution counts within the Planet Nine project

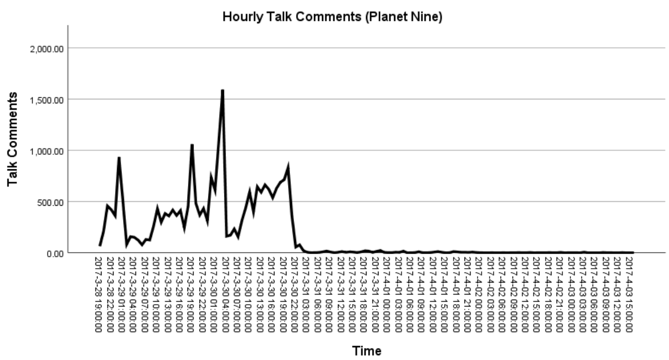


Figure 27: Line graph showing hourly talk comment counts within the Planet Nine project

rapid and greater rate than in Exoplanet Explorers, falling to just 10,206 classifications. Unlike Exoplanet Explorers, however, the classification rate in Planet Nine gradually increased over the following day - with a second, smaller peak not coinciding with the second broadcast event, but in fact falling some 6 hours later. Once again, the classification rates fell rapidly, only to progressively rise at a slower rate on the third and final day of the broadcast events, reaching the smallest peak shortly after the third broadcast. Finally, classification rates fell once again to just hundreds of classifications an hour and despite some minor variation, did not recover after the broadcast cycle concluded.

On the whole, talk comment rates within Planet Nine follow task contribution rates, with some significant differences. Firstly, initially talk comments rise at a slightly faster pace than task contributions, with a peak coinciding with the initial broadcast event and a second, greater peak in the following hours. A similar phenomenon occurs with the second event, with a peak during the broadcast itself - a peak not seen within the task contribution rates - followed by a second much greater peak in talk contributions in the hours after the broadcast, which occurs *before* the peak in task contributions. This is also seen on the final day of broadcasting, with an initial peak in talk contributions during the broadcast, followed by a peak in task contributions after the broadcast has concluded. As with Exoplanet Explorers, talk contribution rates are more prone to variation, but it is the second broadcast event and the hours following which generates the greatest number of talk comments from participants. As with talk contributions, participation rates fall and do not recover following the final broadcast, but there are noticeably larger fluctuations in talk comment rates, although it should be highlighted that these represent between just 1 and 11 hourly comments, rather than the hundreds or thousands of hourly talk classifications during this period.

8.4 Correlation Between Talk and Task

When considering hourly contribution rates, both projects demonstrate strong correlation between task and talk contributions. In Planet Nine, talk activity appears to be inherently linked with task workflows, with contribution rates in both diminishing at a similar rate and at similar times upon the completion of the BBC Stargazing Live broadcast cycle. Due to the more sudden conclusion of Exoplanet Explorers, it is more difficult to state with certainty whether a

Table 19: Correlation coefficient Results for Exoplanet Explorers and Planet Nine, including the coefficient τ, the p-value *p* and the sample size *N*

Name	τ	***P***	***N***
Exoplanet Explorers (Hourly)	0.70	0.000	95
Exoplanet Explorers (All Volunteers)	0.22	0.000	8164
Exoplanet Explorers (Only Talk)	0.294	0.000	1155
Planet Nine (Hourly)	0.70	0.000	140
Planet Nine (All Volunteers)	0.21	0.000	20346
Planet Nine (Only Talk)	0.30	0.000	2526

similar phenomenon would have occurred, but there is a notable reduction in both talk and task contribution rates – once again at a similar rate and time – shortly prior to the closure of the project.

Although in both projects participant contributions appear to be linked with broadcast events, interestingly participants continue to participate after the final broadcast event – that is, after the announcement of the final project results. Unlike other Zooniverse projects, the time limitation and desire to announce discoveries live on air results in a more explicit link between talk and task within these projects than in activities such as Galaxy Zoo. For example, it is possible (though open to debate) that scientists would have found the Hanny's Voorwerp eventually and without such a discovery, Galaxy Zoo still achieved its goal of morphological galaxy classification. In contrast, within Exoplanet Explorers and Planet Nine, there existed a real-time element, with the desire to make announcements about project findings during Stargazing Live broadcasts. Talk comments appear to drop off at about the time of the final broadcast, suggesting that it is calling attention to these potential discoveries that predominantly motivates talk activity.

On the other hand, while both projects demonstrate statistically significant correlations between individual participants' talk and task contributions, these correlations are low enough to be considered negligible. As figures 28 and 29 show, the majority of participants make very few talk comments in both projects and a very small number of participants make a disproportionate number of talk comments – particularly in the case of Planet Nine. The participant responsible for the most tasks in Planet Nine made just three talk comments although the participant responsible for the second largest number of classifica-

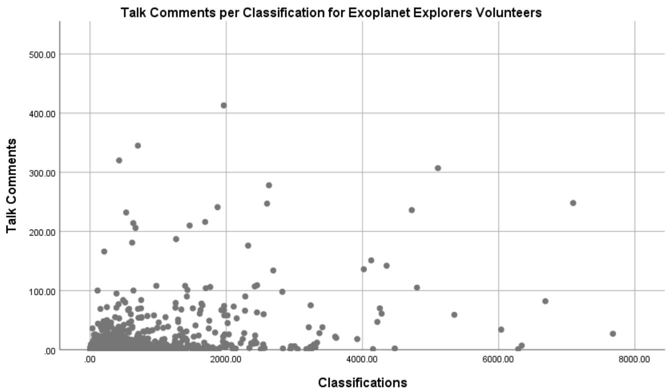


Figure 28: Scatter graph showing classification counts and talk comments for all Exoplanet Explorers volunteers who engaged with talk.

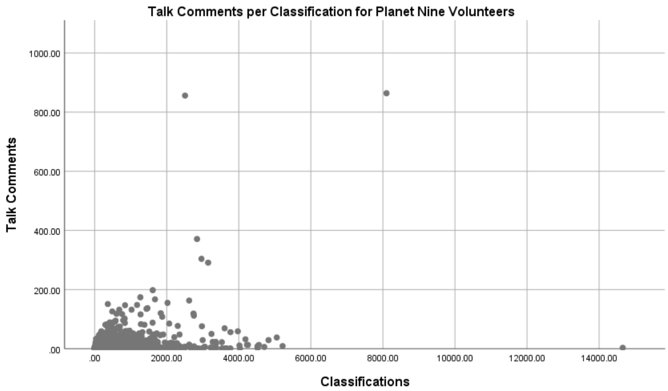


Figure 29: Scatter graph showing classification counts and talk comments for all Planet Nine volunteers who engaged with talk.

tions also made the largest number of talk comments. In Exoplanet Explorers, similarly, the highest contributing participant to engage with talk made just 27 talk comments, while the most active talk participant made around 2000 classifications – an impressive figure, but one that is significantly less than the highest contributing volunteers.

8.5 Distribution of Effort

The distribution of effort in both projects is highly skewed, with a small proportion of participants contributing most of the work. This can be seen in the Gini coefficients of 0.71 for both Planet Nine and Exoplanet Explorers. These figures are lower than the average findings detailed by Cox et al. (2015b) and by Sauermann and Franzoni (2015) of 0.91 and 0.81 respectively. The values provided by Cox et al., and Sauermann and Franzoni, however, were calculated

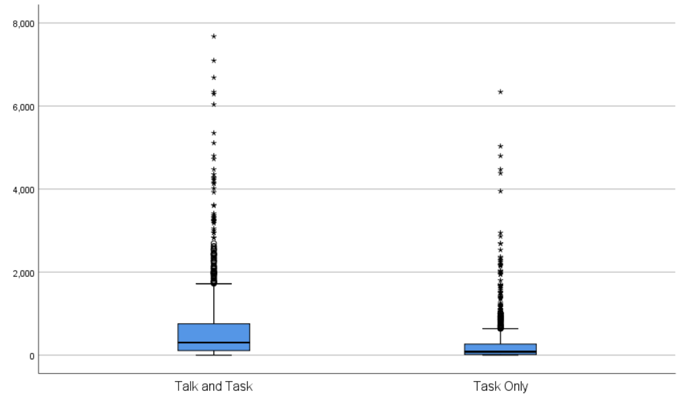


Figure 30: Box and whisker plots for Exoplanet Explorers showing the range of classifications received for task only and talk and task contributing volunteers

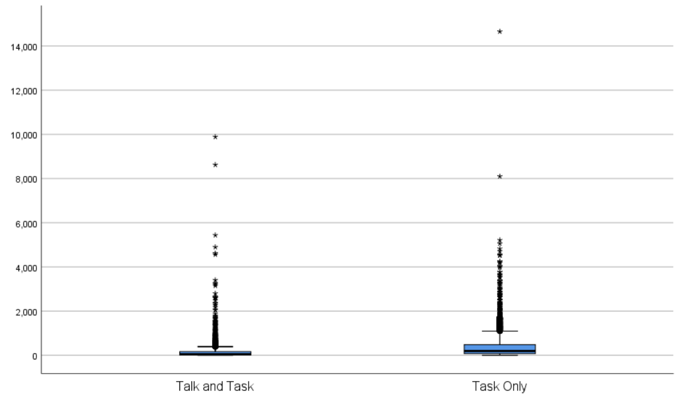


Figure 31: Box and whisker plots for Planet Nine showing the range of classifications received for task only and talk and task contributing volunteers

prior to the introduction of the Panoptes system and the resulting rise in the availability and number of Zooniverse projects. Since no more recent figures have been published, it is not possible to conclude with any certainty whether the distribution of effort in both projects is truly below average.

Contrary to the findings in many other projects, when accounting for outliers, a box and whisker plot of Planet Nine participants who contributed only to the task element of the project exhibits a higher lower and upper quartile and mean number of contributions than the talk and task group, as shown by figure 31. Even among those participants considered outliers, the maximum number of contributions was made by a participant who did not make any talk comments, but instead made over 14,000 classifications - far in excess of the highest contributing talk participant. The second highest contributing task participant also outperformed many of the talk contributors, with approximately 8,000 classifications.

The opposite is however true in the case of Exoplanet Explorers, which generated fewer classifications overall. The most highly performing contributors were all active in both talk and task, with approximately 7,000 classifications, as shown by figure 30. Similarly, when accounting for outliers, the lower and upper quartile and mean are all significantly higher among the talk and task group than in the task only group. Nevertheless, even among those participants who contributed solely to the task element, classification counts are relatively high with many participants contributing thousands of classifications. Certainly, however, both projects provide evidence to question the view that the most active talk participants are always the most active task participants.

8.6 Hypothesis Testing

This section outlines the hypothesis formation and testing process for each of the two projects. It should be noted that the specific hypotheses to be tested in each project are the same, on the basis of similar literature findings and the similar nature of each project in terms of aim, characteristics and overall project lifespan.

8.6.1 Hypothesis Formation

As in chapter 7, hypothesis testing is necessary to better understand the association between task completion and sociality within the analysed projects. However, unlike in the case of chapter 7, there were no interviews from which to gauge the views or assumptions of the EyeWire team on whether any such relationship might exist. While two informal calls were held, the team made it clear that they had done little analysis of the data and did not have strong opinions about any possible link. Nonetheless, one member of the team viewed the Exoplanet Explorers project as more effective, due to the level of dialogue that the team had established with the community.

On this basis, I turned to the literature to shape and formulate the hypotheses to be tested. In terms of a link between task completion and talk comments, the literature findings are somewhat varied, as demonstrated in chapter 4. On this basis, I once again chose not to give greater weight to the null or non-null hypothesis and instead both are presented here without making value judgements as to which is most likely to be correct.

Firstly, the literature suggests that there may be a link between task submission and task completion (Luczak-Roesch et al., 2014; Tinati et al., 2014). Although there is a potential relationship between the length of time a user has contributed to a project before he/she makes his/her first talk comment and the resulting number of classifications that user makes ((Jackson et al., 2016a), given the fact that both projects were new when launched and lasted only a brief time, there was insufficient data to analyse such a relationship. Similar findings regarding the average active lifespan of a project as suggested by Jackson et al. (2014) and by Curtis (2018a) were also incompatible with the brief lifespans associated with the projects examined.

In less concrete terms, Luczak-Roesch et al. (2014) suggest that there may be a relationship between whether a participant receives a reply or other form of acknowledgement of his/her talk activity and resulting classification counts. There is no quantitative evidence of this within the literature, but qualitative and anecdotal findings reported by Curtis (2018a) and by Darch (2017) support the testing of this hypothesis and further raise the possibility that participants are further influenced by whether the respondent is a project scientist, or as reported by Kraut et al. (2012), is otherwise a figure of some importance within the community.

On this basis, for each of the two projects, the following hypotheses were formulated and examined:

$H_E and H_F 1$ Engagement in Talk is associated with a significantly higher number of task contributions.

$H_{E-null} and H_{F-null} 1$ There is no significant association between engagement in Talk and the number of task contributions made.

$H_E and H_F 2$ Receiving at least one reply to a Talk comment is associated with a significantly higher number of task contributions than not receiving a reply.

$H_{E-null} and H_{F-null} 2$ There is no significant association between the number of task contributions made and receiving at least one response to a Talk comment.

$H_E and H_F 3$ Receiving at least one reply to a Talk comment is associated with a significantly higher number of Talk comments than not receiving a reply.

$H_{E-null} and H_{F-null} 3$ There is no significant association between the number of Talk comments made and receiving at least one response to a Talk comment

$H_E and H_F 4$ Receiving at least one reply from a project scientist is associated with a significantly higher number of task contributions than receiving only replies from other participants.

$H_{E-null} and H_{F-null} 4$ There is no significant association between the identity of the respondent and the number of task contributions made.

$H_E and H_F 5$ Receiving at least one reply from a project scientist is associated with a significantly higher number of Talk comments than receiving only replies from other participants.

$H_{E-null} and H_{F-null} 5$ There is no significant association between the identity of the respondent and the number of talk comments made.

Table 20: Summary of hypothesis, *H* Mann-Whitney U test outcomes for Exoplanet Explorers. Z-crit for $p<0.01$ = -2.33

H	Condition	Z-Stat	*d*	Conclusion
H_E 1	Number of Task Submissions (Talk and Task vs Task Only - Exoplanet Explorers)	-20.80	0.68	Hypothesis confirmed. Talk participants contributed more tasks on average than talk participants ($p<0.01$, medium effect size)
H_E 2	Number of Task Submissions (Received Reply vs No Reply)	9.3	0.55	Hypothesis confirmed. Participants who received a reply to their talk comments contribute more tasks on average than those who did not ($p<0.01$, medium effect size)
H_E 3	Number of Talk Comments (Received Reply vs No Reply)	-0.47	N/A	Hypothesis rejected. There is no statistically significant difference in the number of talk comments made by those who received at least one reply to a talk comment and those who did not (p=0.64)
H_E 4	Number of Task Submissions (Reply from Scientist vs Other Participants)	0	N/A	Hypothesis rejected. There is no statistically significant difference between the number of tasks completed by those who received at least one reply to a talk comment from the science team and those who only received responses from other participants (p=1)
H_E 5	Number of Talk Comments (Reply from Scientist vs Other Participants)	-8.18	0.56	Hypothesis Confirmed. Talk participants who received at least one reply from a science team member made more talk comments on average than those who only received responses from other participants ($p<0.01$, medium effect size)

Table 21: Summary of hypothesis, *H* Mann-Whitney U test outcomes for Planet Nine. Z-crit for $p<0.01$ = -2.33

H	Condition	Z-Stat	*d*	Conclusion
H_F 1	Number of Task Submissions (Talk and Task vs Task Only)	-35.44	0.51	Hypothesis confirmed. Talk participants contributed more tasks on average than talk participants ($p<0.01$, medium effect size)
H_F 2	Number of Task Submissions (Received Reply vs No Reply)	12.46	0.5	Hypothesis confirmed. Participants who received a reply to their talk comments contribute more tasks on average than those who did not ($p<0.01$, medium effect size)
H_F 3	Number of Talk Comments (Received Reply vs No Reply)	15.81	0.64	Hypothesis confirmed. Talk participants who received at least one reply to a talk comment made more comments on average than those who did not ($p<0.0.01$, medium effect size)
H_F 4	Number of Task Submissions (Reply from Scientist vs Other Participants)	-6.69	0.30	Hypothesis confirmed. Talk participants who received at least one reply from a science team member made more talk comments on average than those who only received responses from other participants. ($p<0.01$, small effect size)
H_F 5	Number of Talk Comments (Reply from Scientist vs Other Participants)	-14.59	0.51	Hypothesis Confirmed. Talk participants who received at least one reply from a science team member made more talk comments on average than those who only received responses from other participants ($p<0.01$, medium effect size)

8.6.2 Hypothesis Testing Results

In both projects, participants who engage with talk complete significantly more classifications than those who only engage with the task elements of each project. Nevertheless, this effect was greater in Exoplanet Explorers (d=0.68) than in Planet Nine (d=0.51). Clearly then, while the interquartile range, lower and upper quartile and top performing participants were all higher among task-only contributors in Planet Nine, the overall average classification rate per participant is still lower among this group - likely due to the sheer size and the large quantity of low-performing participants, which represent a significantly higher proportion of the total number of contributors than in the talk and task contributor group.

There also appears to be a truly social element to the interaction between talk activity and task completion. Participants who received at least one response to their talk comments and queries completed more task submissions on average than those who engaged in talk without receiving a reply. In the case of Planet Nine, a similar effect was seen in terms of the number of talk comments that those who received replies left within the project, with high statistical significance ($p < 0.01$) and a medium effect size, but no such effect could be confirmed within Exoplanet Explorers. This appears to extend to comments and interaction from members of the project science team. In Exoplanet Explorers, there was no statistically significant difference in the number of tasks completed by volunteers who interacted with the science team and those who only interacted with other volunteers, but a significant effect in terms of talk comments left by those receiving replies from science team members - with a medium effect size (d=0.56). In Planet Nine, however, a significant effect *was* observed in the number of tasks completed by those who interacted with the science team, with a small effect size, suggesting that under the right circumstances, feedback from the science team may spur increased action in volunteers. As with Exoplanet Explorers, a medium sized effect was also observed in the number of talk comments left by volunteers interacting with the science team. It should be reiterated, however, that correlation does not imply causation and it is not possible to state conclusively that replies from science team members influence participant activity.

8.7 Discussion

On the one hand, hourly talk and task contributions are strongly correlated in both projects. As task classifications increase, so too do talk comments and at the end of the course of both projects, task and talk contribution rates sharply decrease at a similar time, to a similar extent. Similarly, in both projects the broadcast events lead to a significant increase in both task and talk activity. Despite the limited time available to participants, volunteers in both projects who contribute to talk complete far more classifications on average than participants who only complete microtasks. Even so, while talk participants made significantly more classifications in both projects, non-talk participants in both groups were able to contribute many thousands of classifications in the limited time available, raising questions about the extent to which there is a link between productivity and sociality. In both projects, the task provided to volunteers was extremely simple – a binary yes or no choice – and there was no other task option to distract from the repetitive workflow, a key opportunity previously ascribed to Talk (Jackson et al., 2014). Similarly, while talk has been described as essential for the discovery process (for example, by Tinati et al. (2015b), thousands of task participants completed thousands of classifications without making a single talk comment to draw attention to a potential finding. Assets within Zooniverse projects are of course served randomly and so it is possible that participants simply did not find anything, but even so, the addition of deadlines and the opportunity for live feedback on results had no effect on the majority of participants. Certainly, however, under the right conditions – for example, in the short term and with clear goals – participants are able to complete large numbers of task classifications without the need to engage in other activities or the drive to leave a project.

Furthermore, for individual participants, there is little correlation between task and talk activity. The specific correlations displayed in the project are very similar to the correlation between the number of tasks completed and chat messages sent in EyeWire, as calculated by Tinati et al. (2015a), where Kendall's τ of 0.3 is roughly equivalent to Spearman's ρ of 0.43 (Gilpin, 1993). More specific to the Zooniverse, Tinati et al. (2014) describe how task contributions in Zooniverse significantly outnumber talk contributions, a factor also observed in Exoplanet Explorers and Planet Nine. However, it appears the correlation between individual talk and task contributions is weaker in these Stargazing

Live projects than in the projects studied by Tinati et al. (2014), as while no correlation coefficient is given, the graphical representation of the correlation between talk and task contributions is clearly stronger than the weak to negligible correlation of 0.3 described here.

While engagement with talk is associated with higher classification levels overall, these are highly variable and the number of talk comments made by a volunteer is no indication of how engaged he or she is in the project and how many tasks he or she has completed. Darch (2017) and Heaton and Torres (2015) suggest that participants expect scientific results within relatively short time-frames and that in the absence of these results participants can be demoralised and demotivated and this was supported by participant one during the interview phase, albeit in the context of EyeWire. Even so and despite the popularity of scientific research as a factor influencing intrinsic motivations, the opportunity to hear about scientific findings in a relatively short time-frame has little impact on participant engagement patterns. The vast majority of participants still contribute very little and both talk and task activity are reliant on a small number of highly active users, even when coupled with a final deadline – a factor which Kraut et al. (2012) suggests should encourage contributions. While the distribution of effort is somewhat greater in both Exoplanet Explorers and Planet Nine, it is not possible to say with certainty whether this is a result of increased engagement, or simply that the project lifespan was too short for top performing participants to build up a greater number of contributions.

While the previous literature has predominantly associated task completion with talk contribution on at least some level, the findings from this chapter raise questions about this. In fact, even among those participants who actively contribute to talk, simply posting comments is not the greatest indicator of task contribution levels. In both projects, receiving a reply or comment further boosted the number of contributions left by participants, with a further boost in Planet Nine when the reply was made by an individual flagged as a project scientist. This suggests that it is not merely exercising new skills, autonomy or relief from repetition that volunteers derive from talk, but that there *may be an interaction- and sociality-based component in the factors that govern individual volunteers' productivity.* At the same time, Luczak-Roesch et al. (2014) identified poor response rates and times within project Talk platforms, noting that volunteers could expect to wait hours or even days for a response.

The significant and sudden impact on task and talk contributions observed within each project is similar to the effects of radio broadcasts observed in Galaxy Zoo. In a manner reflective of Exoplanet Explorers, the first days of Galaxy Zoo generated little interest from volunteers, until a radio channel broadcast interviews with project scientists, leading to a sudden, but lasting increase in the number of participants and contributions to the project (Morais et al., 2013). Ensuring potential participants hear about a project, then, represents a significant barrier to participant recruitment and resulting project productivity. This in itself is not a new finding, as Iriberri and Leroy (2009) and Kraut et al. (2012) both discuss the difficulties associated with participant recruitment to online communities and the issues with relying on word-of-mouth campaigns. Yet it is still odd to note that participant recruitment poses such an issue in the Zooniverse. Strong cross project migration has been observed between projects with similar goals and the Planet Hunters project – with identical goals and an extremely similar task to Exoplanet Explorers is one of the most successful in Zooniverse history, having received contributions from over 160,000 volunteers and at one time being responsible for 25% of all tasks completed within the platform (Curtis, 2018a).

Even so, these findings, in line with findings from Luczak-Roesch et al. (2014), clearly demonstrate that task engagement is not necessarily an indicator of talk engagement. Projects which are appealing in terms of talk – like Exoplanet Explorers or Seafloor Explorer as detailed by Luczak-Roesch et al., – are not necessarily the most appealing projects in terms of task completion. In both of these examples, the number of talk participants outnumbers similar projects, but the number of task participants is significantly lower. Talk nevertheless plays an important role in participant motivations as an opportunity for feedback and learning. Where this feedback is provided, participant motivations are strengthened (as suggested by Jennett et al. (2016)) and this enhances participant productivity. As Jackson et al. (2014) and Morais et al. (2013) suggest, there is also a link between talk activity and long-term contribution patterns, although the projects discussed here are inadequate to observe such a phenomenon. This suggests that while participants in both projects were able to contribute significant numbers of contributions in the short term, such participation would not continue and these initially heavy contributors would be overtaken by less active contributors who remain attached to projects for longer.

8.8 Summary and Conclusions

This chapter has described task and discussion activity within two 48 hour projects conducted in conjunction with ABC and BBC Stargazing Live. The key findings of this chapter are as follows:

- Task completion and discussion contribution rates are strongly correlated, even in situations which encourage volunteers to prioritise one action or the other.
- Conversely, the number of tasks completed by participants greatly outnumbers the number of task contributions made.
- Participants who engage in talk make significantly more contributions on average than participants who do not.
- This effect is greater for participants who receive at least one reply to their talk comments and may be greater still among those who receive responses from project scientists – specific findings across the two projects were inconclusive.
- Even so, in the short-term, participants who do not contribute to talk are still capable of completing significant numbers of classifications, potentially outperforming many of the talk participants.

In the final chapter, I will synthesise and discuss the significance of each of the previous chapters, alongside a discussion of the key conclusions and associated limitations of this thesis.

Chapter 9

Discussion and Conclusions

This chapter serves to summarise and discuss the results of the previous chapters in the context of the overall research questions and aims of this thesis. Firstly, a summary is given of the methods and contributions of each of chapters 4 to 8. Findings are then synthesised based on the main topics of each of the 3 research questions outlined in the introduction, with an additional focus on the consideration of the three background theories described in chapter 2 – collective action, distributed intelligence and social machines. Opportunities for future work stemming from and related to the conclusions of this chapter are then presented, along with limitations stemming from the methodological approach used and projects studied. This chapter – and this thesis – finishes with a set of conclusions aimed to draw together each of the main points derived from this research.

9.1 Summary of Contributions

Throughout this thesis, I have presented findings and conclusions from across 5 main studies, focusing on the nature of VCS and the role that sociality plays in its design and efficiency. This section outlines the methodology and data sources used for each chapter, as well as a brief summary of the main findings and themes resulting from each study. These findings are then synthesised and discussed in more detail in the following section.

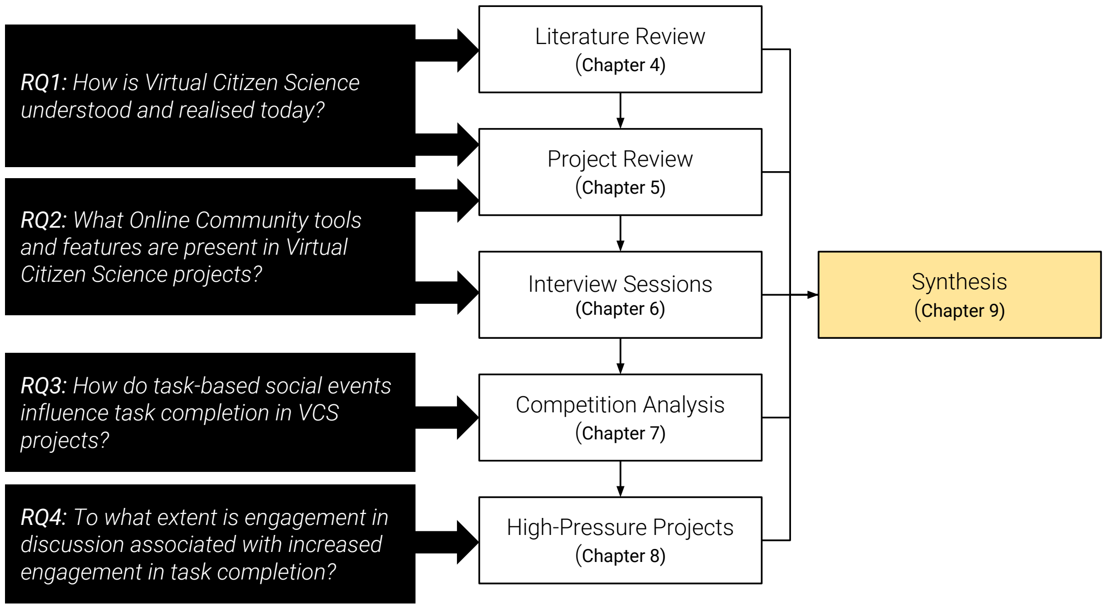


Figure 32: This chapter summarises and integrates the findings of each of the five previous chapters.

9.1.1 Transdisciplinary Literature Review

Chapter 4 described the results of a transdisciplinary literature review drawing on 61 publications from across 5 databases. These results focused on five main areas: the tasks and activities that make up VCS, design challenges and potential solutions within the literature, participant motivations, participant engagement patterns and the form that sociality takes. Based on the 62 projects identified, VCS is predominantly used within astrophysics and biological sciences research, making use of images and simple categorising and classifying tasks. Accuracy and data quality, the quantity of contributions, attracting participants and retaining a sufficiently large community of volunteers make up the main design challenges associated with VCS methodologies. Data quality issues are generally solved through the use of redundancy and aggregation of results, which has a knock-on effect for the efficiency of projects. To motivate contributions, projects may draw on gamification, sociality in the form of competitions and collaborative goals, or design interventions designed to emphasise the significance of projects. Participant recruitment issues were not a prominent focus of the literature, but studies suggest that these too may be solved through design interventions and project framing. Similarly, participant retention, while a significant issue, was not discussed in great detail and at least one source considers high participant turnover rates inevitable (Sauermann and Franzoni, 2015), but there are indications that sociality, collaboration and the nature of assets all play a role in participants' decisions to remain with projects in the long-term.

In terms of motivations, VCS participants are all volunteers and long-term participants report high levels of motivation stemming from altruism and a desire to contribute to science. Intrinsic motivations – that is, participants' own interests in science, tasks or assets – are the greatest factor driving the initial decision to contribute to projects and this motivation persists as participants learn more about the project and potentially earn extrinsic rewards if these are offered. Even so, engagement in VCS is characterised by a large proportion of momentary participants who dabble briefly in projects before leaving and not returning. To achieve project aims, designers are relient on a small number of highly active users who contribute the bulk of contributions.

Social interaction within projects may occur within integrated discussion interfaces, linked forums or custom interfaces or even external services such as IRC and may be synchronous or ascynchronous depending on the service used. Participants interact with one another and less commonly, moderators, administrators and scientists while – or between – enaging in tasks. Sociality is invaluable as an opportunity for participants to learn how to complete tasks, to escape the potentially monotonous task workflows in VCS by exercising new skills and autonomy and also to self-organise around tasks or events. Even so, *lurking* is common and participants may share negative opinions which impact the motivation of their fellow participants to engage with tasks. As a result, while a small number of studies have associated increased task activity with discussion, these findings are ambiguous.

9.1.2 Project Survey

To better understand the specific features present in VCS projects, chapter 5 presented the results of a survey of 48 VCS projects, identifying features through a structured walkthrough process based on 4 themes drawn from the wider literature. These themes included: task visibility, goals, feedback and rewards. Task visibility is low across each of the surveyed projects, with tasks assigned to participants almost exclusively at random and with no option to view the submissions made by other participants or the wider community. Goals occur in approximately half of projects and predominantly take the form of classification challenges, where the community of participants must achieve a certain number of classifications by a given deadline, with aims that could not be achieved by any one individual. Feedback mechanisms were found in a similar number of projects, but were predominantly task-contingent, with participants receiving feedback on the quantity of submissions made, either by individuals or by the community as a whole. Performance-contingent feedback regarding task accuracy was significantly less common and in many cases, offered only through discussion features. Finally, rewards were offered in a large proportion of projects, predominantly in the form of status rewards, offering players indicators that could be displayed to the wider community or in some cases, new activities that other players could not achieve. Extrinsic motivators such as points and rewards were found in approximately a quarter of projects and were predominantly associated with games with a purpose, although some less gamified projects such as Notes from Nature also featured rewards.

9.1.3 Interview Sessions

Chapter 6 detailed comments and findings from three semi-structured interview sessions carried out with the 6 full-time members of the EyeWire team. The main aim of this chapter was to understand the factors influencing the VCS design process, to understand why online community and discussion features are added to VCS projects and to identify how players had made use of these features in the ontext of EyeWire. Several themes arose during the interview process. Participants highlighted the importance of productivity and cost-effectiveness to the perception of EyeWire as a successful project, noticing trade-offs and balances necessitated with task accuracy and participant engagement and enjoyment. With regard to observations of participant motivations, the team noted significant intrinsic motivations among the community, but all participants felt that extrinsic motivators – points, badges and other rewards – were the most effective feature within EyeWire. The issue of altruism was somewhat controversial, particularly in light of the difficulties associated with the Scythe Complete feature, which had initially relied on participant altruism with little success.

In terms of online community features, the team had been willing to introduce greater task visibility, but had unfortunately been unable to due to the simple, repeatable nature of EyeWire tasks. Feedback and rewards were both viewed as very important by the team, who felt that feedback could serve as a reward mechanism, making up for the lack of progression within the game as participants receive messages congratulating them on high performance and constantly strive for self-improvement by observing their own personal accuracy score. Goals, however, were particularly important within EyeWire, having revitalised the project after an initial slump and driving more volunteers to the project, but also increasing the efficiency and productivity of the project and potentially, community health and the accuracy of submissions.

Chat and associated discussion was one of the most important features identified by the team, particularly as an escape from the otherwise repetitive nature of the tracing tasks. Beyond this and given the difficulty of neuron tracing, chat also played an important role in training participants, although this was more true of the earlier days of EyeWire than currently. Although the team were somewhat ambiguous with regard to linking chat activity to productivity, they drew links between specific chat features and player desires to complete tasks. These results expanded on and offered an explanation for some of the observations made within the project review, as well as offering a basis for the competition analysis process.

9.1.4 Competition Analysis

Based on the imortance of competitions within EyeWire as described by the team during the interview process, hypothesis testing was conducted drawing on a year of participant contribution and competition data from the EyeWire project. Major competition events are associated with an increased number of players, even among those players who choose not to participate and are also associated with a significant increase in the number of cubes contributed, although not among individual players who do not compete. Players who choose to engage with competitions are a small, but highly active proportion of the EyeWire player base, contributing significantly more cubes than non-competing players even on days when competitions do not take place. However, while competitions are associated with an increase in discussion, there is no association between competitions and less engaging tasks – specifically scything and completing – in spite of the importance these tasks hold for the efficiency of the project.

9.1.5 High Productivity Project Analysis

Finally, chapter 8 presented results from two high pressure projects conducted over the course of 48 hours, during which time participants had to conduct work much more quickly than in ordinary VCS projects. The findings demonstrate that even with relatively immediate feedback on classifications and explicit requirement for collective action, the majority of participants contribute very little work, suggesting that this may be an insurmountable challenge for projects as long as they draw on volunteer efforts. On the whole, participants who engaged with Talk completed more classifications than volunteers who didn't. At the same time, there was little correlation between the number of tasks completed and the number of talk comments a participant left – that is, the most active task participants are **not** necessarily the most active talk participants. Moreover, in both projects, many of the top contributing participants did not leave talk comments at all.

9.2 Synthesis of Findings

This section synthesises and compares findings from each of the qualitative and quantitative studies detailed within chapters 4 to 8. These are divided into relevant themes relating to the three research questions outlined in the introduction – the nature of VCS, online community features in VCS, sociality and productivity and the theoretical framework, particularly VCS as an example of collective action.

9.2.1 The Nature of VCS

Despite the large crowds of participants involved in VCS, the task elements of projects are not truly social activities. As demonstrated in chapter 5, participants are assigned tasks at random and generally have no option to work on tasks in partnership with other players. On the one hand, VCS draws heavily on redundancy and aggregation (Wiggins et al., 2011) and as suggested by Mugar et al. (2014), it is true that such methods somewhat require tasks to be completed in isolation. The EyeWire team, however, hint at a further reason for this: a lack of opportunities for collaboration. VCS tasks – like other forms of microtask crowdsourcing – are designed to be simple and repeatable, such that they are accessible to a large potential audience, without any prior specialist knowledge (Bu et al., 2016). This simplicity, however, hampers collaboration as there is largely no need for specialised roles for such simple tasks. Further evidence for this can be seen in the fact that the two projects identified within this thesis that *do* allow for collaboration are those with the more complex tasks – FoldIt and EyeWire (Curtis, 2015b; Kim et al., 2014). In both cases, this collaboration is delayed, with later players building on earlier submissions and in each case, the collaboration is available only to a small number of trusted players, be they promoted in the case of EyeWire or team mates in the case of FoldIt. The EyeWire team have expressed the desire to introduce further opportunities for participants to work together as teams or groups, but found that such approaches are incompatible with the nature of VCS tasks.

As a result, sociality and disussion are not required to engage in VCS activities and there is little need for explicit interaction between players. Discussion and interaction can play an important role, however, in facilitating VCS tasks by training players, as in the absence of discussion features, tutorials are rarely completed and in some cases, are insufficient to train participants in how to complete project tasks (Kim et al., 2014; Ponciano et al., 2014). The use of discussion features as an opportunity for learning is widespread in both simple and more complex tasks, having been identified by the EyeWire team as mentioned in chapter 6, but also being described previously in FoldIt and potentially in Planet Hunters and Seafloor Explorer (Curtis, 2015b; Mugar et al., 2014, 2015). Interaction plays a further facilitative role as activities become more complex, with the EyeWire team collaborating with high performing players through the chat interface to perform more complex activities within the project than could otherwise be accomplished alone by either group. This calls to mind the rare but valuable phenomenon of citizen-led serendipitous discoveries within Galaxy Zoo, driven initially by volunteers and later, by interaction between volunteers and project scientists (Tinati et al., 2015b).

9.2.2 Online Community Features in VCS

Arguably the most important use of sociality in VCS is in converting individual workflows into online communities. The project review process found evidence of a wide-ranging selection of features used to attract contributions to VCS projects. Despite the lack of focus on extrinsic factors within the motivations literature, rewards appeared in almost a quarter of all projects and a social title or role designed to be displayed to other players was present in over three quaters of projects. Feedback on one's own and the community's progress was also relatively common, appearing in over half of projects. Yet it is the provision of goals which lends the greatest social experience to projects, allowing participants to work together or compete to achieve more accessible aims than the scientific research purposes attached to projects. These goals serve an important design purpose within gamified projects, as the EyeWire team was able to overcome several stated weaknesses of VCS as a genre through the use of goals, allowing participants to win, earn rewards and feel they had accomplished something simply through the introduction of competitions.

Task visibility features are rare in VCS, with projects assigning tasks to individuals on a predominantly random basis, with little opportunity to share submissions. Exactly why this is appears to depend on the specific project in question. Mugar et al. (2014) states that Planet Hunters and Seafloor Explorers both have tasks designed to specifically restrict task sharing and visibility and this somewhat aligns with reports from Wiggins and Crowston (2011) and Wiggins et al. (2011) that redundancy is common within VCS. Conversely, the EyeWire team described a willingness and even desire to introduce greater levels of collaboration to the task, noting that they were hampered by the lack of opportunities for multiple participants to engage in tasks. Competition events which allow weak collaboration among players for specific tasks nonetheless effectively increase the number of submissions made and the number of contributing players within VCS projects. Given the observations made by Crowston et al. (2018), it is likely that task visibility restrictions will become relaxed as projects move towards more collaborative citizen science formats. Even so, the need for quality assurance processes will continue to restrict the extent to which sharing of submissions is enabled within VCS projects.

Similarly, specific task-contingent feedback is rare within VCS projects. Those projects which introduce such feedback generally do so through the use of simulated or gold-standard assets. Participants who wish to seek feedback about more general contributions must do so through social and discussion features. Even so, feedback and learning are key motivational factors within VCS projects, reported by participants across projects. These in turn align with volunteers' intrinsic motivations, as participants seek to learn more about the domain in which they work or even the project they contribute to, as shown by the experiences of EyeWire team. Feedback about project progress is more nuanced. On the one hand, sources including Kraut et al. (2012) state that feedback relative to project progress can drive participation. Conversely, the EyeWire team choose to limit interaction between project participants and scientists, as scientific results progress slowly. In VCS contexts, then, feedback is not as effective as in other online communities.

The sampled literature dealing with motivations within chapter 4 did not discuss extrinsic factors in great detail, largely because such factors were not reported by participants. Nov et al. (2014) suggested that the low score given by participants to reputation mechanisms was linked to their perception of such features with rewards, although specific evidence was not provided for the claim that participants are not motivated by extrinsic factors. Participants from FoldIt and Galaxy Zoo report little interest in reward and competition mechanisms such as leaderboards (Curtis, 2015a; Darch, 2017), althoug Ponti calls into question this view and suggests that it is the Galaxy Zoo team who resist the use of such mechanisms. Tomnod participants also described little interest in the provision of rewards to high performing players (Baruch et al., 2016). Even so, rewards are relatively common in VCS projects, with the option to earn titles or roles present in over three quarters of projects. The EyeWire team suggest that participants are very appreciative of rewards – be they points, badges or physical prizes, although EyeWire is very much a points-based game. Yet throughout these findings there have been suggestions that participants may view scientific outcomes as a form of reward, being disheartened by the slow pace of publications in EyeWire and unwilling to complete tasks which do not directly lead to project science in EyeWire and Galaxy Zoo (Darch, 2017).

9.2.3 Sociality and Productivity

Although in facilitating activities, discussion undoubtedly at least partially contributes to productivity, it is arguably not essential or integral. The vast majority of participants do not engage with discussion features at all, including both short-term “dabblers” described by Eveleigh et al. (2014), but also the “lurkers” described by the EyeWire team who contribute a low to medium quantity of contributions for a sustained period of time. In the short term, as demonstrated by chapter 7 and 8, participants can contribute heavily to projects with little to no use of discussion features. Even among highly active participants, EyeWire players tend to mute the chat window while completing tasks (Tinati et al., 2015a). Yet sources have repeatedly associated the greatest volumes of task contributions with the most active talk participants – across the Zooniverse, in EyeWire and in Planet Hunters (Jackson et al., 2016b; Tinati et al., 2014, 2015a). As shown by 8, this is not entirely the case, as in the short term, task-only participants contribute as heavily as talk contributors, but on the whole, talk participants *do* make more contributions overall than task participants.

Exactly why this may be is still unclear. Discussion and the community are not a significant motivator for participants, either as self-reported or as demonstrated by behaviours in project experiments (Nov et al., 2011a; Tinati et al., 2017a). This is not simply because the majority of participants avoid discussion, as Raddick et al. (2013) described similarly low levels of motivation as reported by Galaxy Zoo participants who were active in – and recruited for the study through – the project forum (since replaced by *Talk*). Discussion and opportunities for sociality – such as sharing achievements with one another – do interact positively with participants' intrinsic and extrinsic motivations. EyeWire participants use chat to discuss neuroscience, the project and the underlying technology itself, while the team use award ceremonies and chat indicators to spur participants to contribute as they receive recognition from fellow players. Opportunities for feedback and learning are also significant affordances of discussion platforms, further influencing participants' intrinsic motivations (Jennett et al., 2016). This can be seen in the greater number of classifications made by participants who received responses to their questions in Exoplanet Explorers and Planet Nine.

Instead, the use of discussion features may be better viewed as an *indicator* rather than a source of motivation to participate in projects. This aligns with the self-reported motivations discussed within chapter 4 and with the description of "dabbling" behaviours observed by Eveleigh et al. (2014). In contributing to a new VCS project, many participants *dabble*, trying out a project for a brief period to see if it aligns to their interests and needs. Talk engagement is a lengthy process, particularly if participants are awaiting a response or asking a question - at least in the case of Zooniverse's *Talk* (Luczak-Roesch et al., 2014). Dabbling periods are also extremely brief - 90 seconds as described by Tinati et al. (2015b). Jennett et al. (2013) explains that learning about projects - through, for example, discussion features - can strengthen and reinforce participants' intrinsic motivations, but if Eveleigh et al.'s findings are accurate, then these participants are already more motivated than the majority of the community - or at least, than the dabblers.

9.2.4 VCS as Collective Action

But what of VCS as a collective action, community-driven process? On the one hand and to an extent, VCS projects rely on volunteers pulling their weight and doing their bit for the good of the project. There is also evidence within the literature that this also drives participants. Diner et al. (2018) and Laut et al. (2016) both demonstrate that the opportunity to see how many submissions fellow participants are making can influence volunteer contribution rates, with participants making more or fewer contributions in line with the wider community. Similarly, collective motives were the most significant factor influencing participation in the Stardust@Home project, while also being important factors in FoldIt and Galaxy Zoo, as well as an experiment conducted by Rotman et al. (2014) (Nov et al., 2014; Reed et al., 2013). Evidence from the analysis of competition participation in EyeWire supports this hypothesis, as the chance to collaborate with other players sharply drove participation, even during minor competition periods without the allure of narratives and progression experiences. This was particularly the case with the marathon competition class, where participants as a collective must complete as many cubes as possible, as quickly as possible.

Even so, project communities may not exhibit collective action in the strictest sense. While Marwell and Oliver (1993) describe collective action as developing once initiatives reach a *critical mass* after the initial input of altruistic participants, it cannot be said that this is true of VCS. Initial participation is predominantly driven by intrinsic motivations and Rotman et al. (2012) observed altruistic motivations only in long-term project participants. The majority of players contribute a very small number of classifications for brief periods of time (Ponciano et al., 2014). This is true even in high-pressure situations where the time available to make contributions is limited and where projects are particularly reliant on each participant doing their bit. Moreover, evidence from the EyeWire team and from Darch (2017) and Woodcock et al. (2017) calls the idea of altruism in long term participants into question, suggesting that these altruistic motives have limits in terms of the influence they have upon participants. Participants are not truly motivated by altruism, but potentially by a reward-style mechanism where they view scientific results as the reward for and outcome of their actions. Even where these results are delivered more immediately and are more explicitly dependent on participants' short-term efforts, such as in the case of the Stargazing Live projects, this is not enough to drive significant participation from most volunteers. Linked to these contribution patterns, Sauermann and Franzoni (2015) further suggests that far from a *critical mass*, projects are constantly in a state of flux and to an extent, death – no matter how slow – as originating members of the community leave and fail to be replaced by sufficiently engaged newcomers.

It might be argued that collective action carries with it certain assumptions – for example, Marwell et al. (1988) state that collective action is commonly seen to rely on social bonds between participants. Similarly, there is the suggestion made by Van Zomeren et al. (2004) that collective action can be viewed in light of social identity theory, as participants engage in collective action to build their own social identities. In the light of these suggestions, sociality and community would be essential for driving participation. It is, after all, only through sociality that participants in VCS could build an identity and social bonds. The EyeWire team showed an awareness of this idea, in discussing the importance of chat for the project:

> "[Chat and email have] been huge in maintaining that relationship... to create that tone of a community where we make sure things were extremely friendly as possible and we make sure that everyone feels included in the chat... And as far as, you know, that kind of community management goes, that has been our main goal. To kind of foster the community and set the tone, so that more older players who have been around longer can pass that along to newer players and keep it going." (*Participant 2*)

This partially ties into Actor Network Theory and the nature of VCS projects as networks of loosely connected human actors and machine and technology elements. These findings suggest that there is a risk in some cases that projects will lose sight of the more vital and expanded role that human agents can play within a network. In this way, project participants are treated almost like machines, rather than as active influences within the project. As demonstrated by the results of chapters 4 and 5, projects do not offer volunteers significant agency, particularly in terms of opportunities to steer the research agenda within projects or project outputs. Participant 1 was particularly cognisant of this issue, both in the context of EyeWire as a game and as a scientific endeavour:

> "They're not just usernames – they're people. They're real people who are enthusiastic and they care and they want to help. So it's not like trying to take advantage of them. It's trying to invite them in to be a part of the thing that you're trying to do. The knowledge that you're trying to create and the discoveries that you're trying to make." (*Participant 1*)

The value of these social relationships – particularly in terms of collective action – is nevertheless debatable. In the case of EyeWire, building trust with the community and engaging in discussion around the scythe complete feature had no impact on participants' willingness to use the feature. Instead, players are observed to be highly motivated by points and by their reputation. In this way, the social identity that participants build is more reliant on rewards and social indicators than on the actual relationships built with others. Furthermore, at least in the short-term, participants can and do contribute heavily to projects without any social interaction or recognition, as demonstrated by the high performing task-only Stargazing Live participants. There is also no link between the number of task and talk contributions made by participants, either in the projects examined by Tinati et al. (2014), or in the high pressure projects examined in chapter 8. Even so, participants who receive responses and interaction from fellow players and scientists contribute more than those participants who do not.

9.3 Limitations and Future Work

There are, of course, additional considerations and qualitative benefits to the use of community features which were outside of the scope of this thesis. For example, further consideration should be given to the value in building a community of interested participants around scientific and research issues and the importance of discussion platforms (as well as social media) in contributing to the formation of such communities. Moreover, as noted by the EyeWire team, such portals and features offer outlets for interaction between volunteers and scientists that can lead to important insights and improved user experiences. These are interesting directions for future qualitative and quantitative research and

Throughout this thesis, data quality has been a matter of some concern in VCS contexts. During the interview stages, the EyeWire team suggested tradeoffs between the quality and quantity of the work done within EyeWire and Gardiner et al. (2012) suggests that this is a common issue across citizen science. There is, therefore, an area of significant importance and interest, which has not been considered within this thesis: what is the relationship - if any - between social activity and community activities and the accuracy and quality of submissions. This is not, however, a simple question to answer. Tinati et al. (2014), Tinati et al. (2015a) and the EyeWire team all suggest that participants who communicate do more work and while this did not hold true in the case of Exoplanet Explorers and to some extent Planet Nine, participants who communicate do tend to contribute more on average than those who do not. In EyeWire, at least, Kim et al. (2014) suggests a link between the amount of work a participant has done and their overall average, although participant 1 cautioned that this was not always the case:

> "Sometimes you'll open up and a player has done 10,000 cubes and their accuracy is still 65%. That's rare, but sometimes that happens!" (*Participant 1*)

Any attempt to explore a connection between discussion and accuracy will have to correct for and consider the impact that extended participation has on the quality of a players' contributions. Where feedback is given, it is also probable that higher performing players are more motivated to do more work and engage with the project, in line with suggestions by Kraut et al. (2012).

Similarly, there are limitations within the methodologies and data drawn upon throughout this thesis. One such limitation is in the generalisability of these results. The methodology and data used within this thesis are not readily conducive to general findings that can be applied to other VCS projects. In particular, the phenomenological interview process used within chapter 6 draws upon a small number of participants from just one project. Similarly, the high pressure, time-limited projects described in chapter 8 are not typical of VCS projects generally. At the same time, however, these results build upon findings within the previous literature and therefore shed light on the extent to which previous, similarly ungeneralisable results can be applied to different VCS projects and contexts. Applying these findings to additional projects, particularly with different domains and task types is an exciting and important area for future work, particularly as VCS diversifies further and new initiatives arise.

A further limitation is in the volume of data available for analysis, particularly in terms of EyeWire competitions in chapter 7. Specifically, the number of days of chat, completion and scything data was relatively limited at just 90 days, which inevitably restricted the number of competition days represented within the dataset. It is possible, then, that the failure to reject the null hypothesis was a result of the low sample sizes involved, rather than an indicator that there was no statistically significant difference between the level of completion or scything in each sample. Even so, the Mann-Whitney U test is suitable for use with very small samples and the use of Z-scores in particular is suitable for any two samples where $N>8$. Since this requirement was met, I am confident that the datasets were still suitable for testing the hypotheses described. Further exploration of the impact of goals and competitions on additional tasks is a timely area of study nonetheless, as in the months following the sampled period, the EyeWire team have begun to introduce competitions specifically for these activities. Moreover, given the findings suggested by Eveleigh et al. (2013), there are outstanding questions regarding the effectiveness of social competitions within non-gamified projects. What effect do these competitions have on participant contribution levels and how do they interact with participants' motivations?

9.4 Concluding Remarks

Any link between sociality and productivity is nuanced and highly dependent on the individual volunteers and projects involved. Sociality within VCS takes multiple, diverse forms and what works for – or is compatible with – any one project may not function in other projects, or at a given time. Since the launch of Galaxy Zoo, for example, the nature and usage of the project forum has evolved sufficiently to encourage the development of *Talk* and greater integration within the project (Tinati et al., 2015b). Generalising findings from any one project is therefore a difficult challenge and what is true today in – for example – EyeWire may not be true in FoldIt next year. At the same time, there are commonalities within projects – the adoption of online community features such as discussion platforms, rewards, feedback and goals and the lack of sociality within tasks themselves. The purpose of these features is not necessarily the same across projects – for example, Tinati et al. (2015b) describes the introduction of talk as designed as a question answering service, while the EyeWire team use chat to break up monotonous tasks.

Overall, however, while tasks themselves are not social, online community features allow participants to gain an understanding of their own performance relative to others, the progress of the project and in some cases, to work together to achieve specific goals. These features are intended – at least in the case of EyeWire – to increase participant numbers and task completion, while social interaction is an opportunity to provide freedom from repetitive task workflows, to better engage participants and to facilitate task completion. Ultimately while the addition of sociality – be that through social competitions or discussion – to tasks is associated with increased engagement and activity from some participants, such features only engage a small minority of participants. Even so, the impact that these features have on productivity and the overall efficiency of projects should not be underestimated. In a context where projects are highly dependent on just a few, highly active players, with potentially temperamental motivations, any increase in productivity is valuable.

There are, in truth, somewhat social tasks within the sphere of VCS. More recently, for example, data analysis initiatives have arisen that make use of social tasks for the purposes of image classification, fitting the collaborative tagging category described by Dunn and Hedges (2013). Specifically, in the ESP game (Von Ahn and Dabbish, 2004), two players work together to label images by attempting to guess the descriptor that the other player may be using. While this activity is not necessarily Citizen Science in and of itself, similar paradigms have been put to use in the Citizen Science space - for example, in Night Knights[1], where players classify images from the International Space Station according to one of six tags, earning points if their tags match those of the paired player (see figure 33) (Celino et al., 2017). These tasks apply social activities to tasks in a different way to many of the projects outlined in this thesis. Due to the rules and requirements of the game, players are prevented from communicating with one another in any way and the format is open and prone to the replacement of one of the players in the pair with a bot (Robertson et al., 2009). There is little difference between the core game activity, where players contribute synchronously and the asynchronous, non-social VCS activities described within this thesis, other than the quality assurance process and the immediacy of its application. Even so, this is an interesting direction for applying sociality to projects in the future and may resolve some of the issues raised by the EyeWire team regarding the difficulties associated with splitting VCS tasks such that they can be completed by groups or teams of players.

Furthermore, while not of particular significance to the issue of productivity,

[1] `https://www.nightknights.eu`

it is important to also consider that opportunities for sociality can have a positive transformative impact on the nature of VCS projects. Each of the projects identified throughout this thesis has corresponded to the *distributed intelligence* level of the framework described in chapter 2, with the exception of Galaxy Zoo Quench. In all cases, this means that participants are active within project research, making significant contributions in terms of producing research data in the form of classifications and metadata and potentially exploring these data. Yet without sociality – without opportunities for discussion or online community features which offer a window into the work of other volunteers – participants are able to exert little influence within the project and are arguably no more part of the research process than the sensors and tools associated with citizen sensing processes. This lack of agency has pronounced effects not only for participants, but for projects and research as well, harming public dissemination and the subsequent understanding and awareness of key scientific issues, resulting in potential issues in areas such as funding (Couvet et al., 2008; Pandya, 2012).

VCS is also a rapidly evolving space, driven by new technologies and advances in what communities have achieved. Projects such as Galaxy Zoo Quench are increasingly pushing the boundaries of what has been accomplished through VCS, it is inevitable that new tasks, features and methods of interaction will be introduced to projects. As suggested by Crowston et al. (2018), these methods will require greater coordination and communication within the community and with project scientists. Even among current VCS projects, as tasks in EyeWire have become more complex, participants and project team members have turned to the chat system to collaborate, accompishing far more than could otherwise be expected without interaction. The role that discussion and the chance to observe other players can play in facilitating VCS must not be overlooked. While this is particularly valuable in EyeWire and FoldIt, where tutorials are lengthy and tasks are complex, it is equally true of simpler projects as projects prioritise 'real' classifications with direct value to science over simulated or tutorial images, or those designed for quality assurance and collaboration (Darch, 2017; Mao et al., 2013a).

Eventually, the EyeWire team foresee a time when human-in-the-loop processes are no longer required, as artificial intelligence takes over and algorithmic processes become more efficient than human processes ever could be:

> "There's some ineffable, but real point in the future when, like, something like citizen science wouldn't be necessary, at least not for this level of

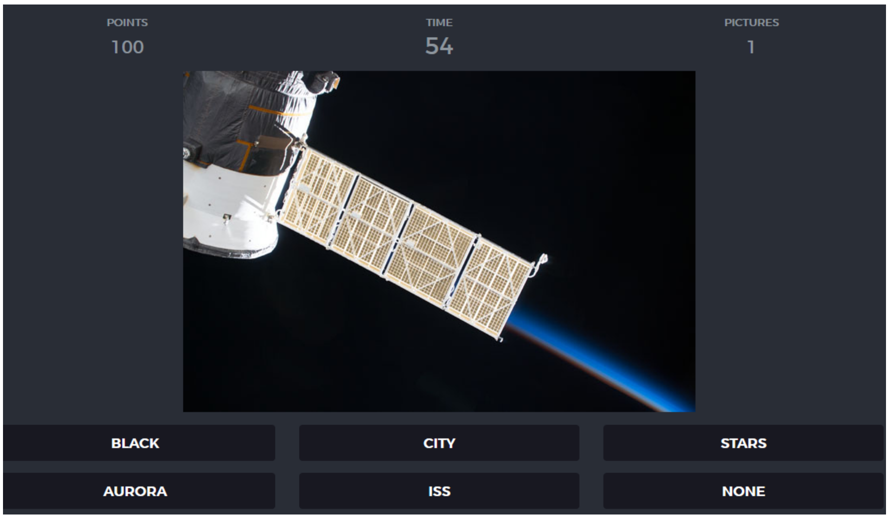


Figure 33: The Night Knights Game With A Purpose interface, which bears similarity to the ESP Game.

> data analysis." (*Participant 6*)

Yet until that day arrives, Virtual Citizen Science is crucial for enabling large-scale analysis of data in a manner that would otherwise be prohibitively slow, expensive or inaccurate – and it is by these benchmarks that it is assessed. VCS has a number of additional benefits – introducing volunteers to science and scientific principles, disseminating results and opening the scientific process. But for *virtual* citizen science, at least, scientific goals are the driving force behind the decision to use human-in-the-loop processes. Therefore, even for those projects that unlike EyeWire are not generally accomplished by teams of paid professionals, ensuring the efficiency of the project is and will continue to be a significant focus for researchers and designers in the years to come:

> "[The goals are] becoming faster at doing it and also kind of getting a good chunk of that done with the game or the project... [In comparison to] if you come out of it without having had any progress on your data that was significant or if you had used just a bunch of undergraduates." (*Participant 5*)

References

Mokhtar Bin Abdullah. On a robust correlation coefficient. *The Statistician*, pages 455–460, 1990.

Allan Afuah and Christopher L Tucci. Crowdsourcing as a solution to distant search. *Academy of Management Review*, 37(3):355–375, 2012.

EC Alexopoulos. Introduction to multivariate regression analysis. *Hippokratia*, 14 (Suppl 1):23, 2010.

Marcio Antelio, Maria Gilda P Esteves, Daniel Schneider, and Jano Moreira de Souza. Qualitocracy: A data quality collaborative framework applied to citizen science. In *Systems, Man, and Cybernetics (SMC), 2012 IEEE International Conference on*, pages 931–936. IEEE, 2012.

Maria Aristeidou, Eileen Scanlon, and Mike Sharples. Profiles of engagement in online communities of citizen science participation. *Computers in Human Behavior*, 74:246–256, September 2017. ISSN 0747-5632.

Arthur Armstrong and John Hagel. The real value of online communities. *Knowledge and communities*, 74(3):85–95, 2000.

K Aye, Megan E Schwamb, Ganna Portyankina, Candice J Hansen, Adam McMaster, Grant RM Miller, Brian Carstensen, Christopher Snyder, Michael Parrish, Stuart Lynn, et al. Planet four: Probing seasonal winds on mars by mapping the southern polar co _2 jet deposits. *arXiv preprint arXiv:1803.10341 (Under Review)*, 2018.

Roger Bakeman. Recommended effect size statistics for repeated measures designs. *Behavior research methods*, 37(3):379–384, 2005.

Avinoam Baruch, Andrew May, and Dapeng Yu. The motivations, enablers and barriers for voluntary participation in an online crowdsourcing platform. *Computers in Human Behavior*, 64:923–931, November 2016. ISSN 0747-5632.

Aaron Bauer and Zoran Popovic. Collaborative problem solving in an open-ended scientific discovery game. *PACMHCI*, 1(CSCW):22–1, 2017.

Patricia Bazeley. *Integrating analyses in mixed methods research*. Sage, 2017.

Nicolas Bencherki. Actor–network theory. *The International Encyclopedia of Organizational Communication*, pages 1–13, 2017.

Tim Berners-Lee. *Weaving the Web: The Past, Present and Future of the World Wide Web by Its Inventor*. Orion Business Books, 1999. ISBN 0752820907.

Gert Biesta. Pragmatism and the philosophical foundations of mixed methods research. *Sage handbook of mixed methods in social and behavioral research*, 2: 95–118, 2010.

Rick Bonney, Caren B Cooper, Janis Dickinson, Steve Kelling, Tina Phillips, Kenneth V Rosenberg, and Jennifer Shirk. Citizen science: a developing tool for expanding science knowledge and scientific literacy. *BioScience*, 59(11):977–984, 2009.

S Borrett and L Hughes. Reporting methods for processing and analysis of data from serial block face scanning electron microscopy. *Journal of microscopy*, 263 (1):3–9, 2016.

Anne Bowser, Derek Hansen, Yurong He, Carol Boston, Matthew Reid, Logan Gunnell, and Jennifer Preece. Using gamification to inspire new citizen science volunteers. In *Proceedings of the first international conference on gameful design, research, and applications*, pages 18–25. ACM, 2013.

Alex Bowyer, Chris Lintott, Greg Hines, Campbell Allan, and Ed Paget. Panoptes, a project building tool for citizen science. In *Proceedings of the AAAI Conference on Human Computation and Crowdsourcing (HCOMP'15). AAAI, San Diego, CA, USA*, pages 1–2, 2015.

Daren C Brabham. *Crowdsourcing*. MIT Press, 2013.

Robin Brouwer. When competition is the loser: the indirect effect of intra-team competition on team performance through task complexity, team conflict and psychological safety. In *System Sciences (HICSS), 2016 49th Hawaii International Conference on*, pages 1348–1357. IEEE, 2016.

Randal Bryant, Randy H Katz, and Edward D Lazowska. Big-data computing: creating revolutionary breakthroughs in commerce, science and society, 2008.

Qiong Bu, Elena Simperl, Sergej Zerr, and Yunjia Li. Using microtasks to crowdsource dbpedia entity classification: A study in workflow design. *Semantic Web*, (Preprint):1–18, 2016.

Vanilson Buregio, Silvio Meira, and Nelson Rosa. Social machines: a unified paradigm to describe social web-oriented systems. In *Proceedings of the 22nd International Conference on World Wide Web*, pages 885–890. ACM, 2013.

Francesco Cappa, Jeffrey Laut, Oded Nov, Luca Giustiniano, and Maurizio Porfiri. Activating social strategies: Face-to-face interaction in technology-mediated citizen science. *Journal of environmental management*, 182:374–384, 2016.

Carolin Cardamone, Kevin Schawinski, Marc Sarzi, Steven P Bamford, Nicola Bennert, CM Urry, Chris Lintott, William C Keel, John Parejko, Robert C Nichol, et al. Galaxy zoo green peas: discovery of a class of compact extremely star-forming galaxies. *Monthly Notices of the Royal Astronomical Society*, 399(3): 1191–1205, 2009.

Christina L Catlin-Groves. The citizen science landscape: from volunteers to citizen sensors and beyond. *International Journal of Zoology*, 2012, 2012.

Irene Celino, Oscar Corcho, Franz Hölker, and Elena Simperl. Citizen science: Design and engagement (dagstuhl seminar 17272). In *Dagstuhl Reports*, volume 7. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2018.

Irene Celino, Andrea Fiano, and Gloria Re Calegari. D4.10 games release - final. 6 2017.

Justin Cheng, Caroline Lo, and Jure Leskovec. Predicting intent using activity logs: How goal specificity and temporal range affect user behavior. In *Proceedings of the 26th International Conference on World Wide Web Companion*, pages 593–601. International World Wide Web Conferences Steering Committee, 2017.

Shu-Chen Cheng, Gwo-Jen Hwang, and Chiu-Lin Lai. Effects of the group leadership promotion approach on students' higher order thinking awareness and online interactive behavioral patterns in a blended learning environment. *Interactive Learning Environments*, pages 1–18, 2019.

Cleo H Cherryholmes. Notes on pragmatism and scientific realism. *Educational researcher*, 21(6):13–17, 1992.

Jacob Cohen. Statistical power analysis for the behavioural sciences hillsdale. *Lawrence Earlbaum Associates*, 2, 1988.

Jacob Cohen. Things i have learned (so far). *American psychologist*, 45(12):1304, 1990.

PF Collaizzi. Existential phenomenological alternatives for psychology, 1978.

MN Converse. Philosophy of phenomenology: How understanding aids research. *Nurse Researcher (through 2013)*, 20(1):28, 2012.

Hugh Coolican. *Research methods and statistics in psychology*. Psychology Press, 2017.

Denis Couvet, Frédéric Jiguet, Romain Julliard, Harold Levrel, and Anne Teyssedre. Enhancing citizen contributions to biodiversity science and public policy. *Interdisciplinary science reviews*, 33(1):95–103, 2008.

Anna Cox, Paul Cairns, Pari Shah, and Michael Carroll. Not doing but thinking: the role of challenge in the gaming experience. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 79–88. ACM, 2012.

Joe Cox, Eun Young Oh, Brooke Simmons, Gary Graham, Anita Greenhill, Chris Lintott, Karen Masters, et al. Doing good online: An investigation into the characteristics and motivations of digital volunteers. In *Leeds University Business School Working Paper*, 16-08. Leeds University Business School, 2015a.

Joe Cox, Eun Young Oh, Brooke Simmons, Chris Lintott, Karen Masters, Gary Graham, Anita Greenhill, and Kate Holmes. How is success defined and measured in online citizen science? a case study of zooniverse projects. *Computing in science and engineering*, 2015b.

Joe Cox, Eun Young Oh, Brooke Simmons, Chris Lintott, Karen Masters, Anita Greenhill, Gary Graham, and Kate Holmes. Defining and measuring success in online citizen science: A case study of zooniverse projects. *Computing in Science & Engineering*, 17(4):28–41, 2015c.

Alycia Crall, Margaret Kosmala, Rebecca Cheng, Jonathan Brier, Darlene Cavalier, Sandra Henderson, and Andrew D Richardson. Volunteer recruitment and retention in online citizen science projects using marketing strategies: Lessons from season spotter'. *JCOM: Journal of Science Communication*, 16(1):A1, 2017. ISSN 1824-2049.

John Cresswell. *Research design*. Sage Publications, 2014.

JW Cresswell and VL Plano Clark. Designing and conducting mixed method research. 3rd ed., sage. *Thousand Oaks, CA*, 2018.

Kevin Crowston and Isabelle Fagnot. The motivational arc of massive virtual collaboration. In *Proceedings of the IFIP WG*, volume 9, 2008.

Kevin Crowston, Erica Mitchell, and Carsten Østerlund. Coordinating advanced crowd work: Extending citizen science. In *Proceedings of the 51st Hawaii International Conference on System Sciences*, 2018.

Vickie Curtis. Online citizen science games: Opportunities for the biological sciences. *Applied & Translational Genomics*, 3(4):90–94, 2014.

Vickie Curtis. Motivation to participate in an online citizen science game a study of foldit. *Science Communication*, 37(6):723–746, 2015a.

Vickie Curtis. *Online citizen science projects: an exploration of motivation, contribution and participation.* PhD thesis, The Open University, 2015b.

Vickie Curtis. Exploring online citizen science in depth: A tale of three projects. In *Online Citizen Science and the Widening of Academia*, pages 19–43. Springer, 2018a.

Vickie Curtis. *Online Citizen Science and the Widening of Academia: Distributed Engagement with Research and Knowledge Production*. Springer, 2018b.

Vickie Curtis. Participant interaction: From online forums to virtual communities of practice. In *Online Citizen Science and the Widening of Academia*, pages 143–165. Springer, 2018c.

Peter T Darch. When scientists become social scientists: How citizen science projects learn about volunteers. *International Journal of Digital Curation*, 12(2): 61–75, 2017.

Richard B Darlington and Andrew F Hayes. *Regression analysis and linear models: Concepts, applications, and implementation*. Guilford Publications, 2016.

David De Roure, Clare Hooper, Kevin Page, Ségolène Tarte, and Pip Willcox. Observing social machines part 2: How to observe? In *Proceedings of the ACM Web Science Conference*, page 13. ACM, 2015.

Marie Delacre, Daniël Lakens, and Christophe Leys. Why psychologists should by default use welch's t-test instead of student's t-test. *International Review of Social Psychology*, 30(1), 2017.

Kristen Dergousoff and Regan L Mandryk. Mobile gamification for crowdsourcing data collection: Leveraging the freemium model. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, pages 1065–1074. ACM, 2015.

Travis Desell, Kyle Goehner, Alicia Andes, Rebecca Eckroad, and Susan Ellis-Felege. On the effectiveness of crowd sourcing avian nesting video analysis at wildlife home. *Procedia Computer Science*, 51:384–393, 2015.

Sebastian Deterding, Dan Dixon, Rilla Khaled, and Lennart Nacke. From game design elements to gamefulness: defining gamification. In *Proceedings of the 15th international academic MindTrek conference: Envisioning future media environments*, pages 9–15. ACM, 2011.

Oxford English Dictionary. Oed online. 2017. *Medium, n. and adj*, 2017.

David Diner, Shinnosuke Nakayama, Oded Nov, and Maurizio Porfiri. Social signals as design interventions for enhancing citizen science contributions. *Information, Communication & Society*, 21(4):594–611, 2018.

David B Duncan. t tests and intervals for comparisons suggested by the data. *Biometrics*, pages 339–359, 1975.

Stuart Dunn and Mark Hedges. Crowd-sourcing as a component of humanities research infrastructures. *International Journal of Humanities and Arts Computing*, 7(1-2):147–169, 2013.

Melissa Eitzel, Jessica Cappadonna, Chris Santos-Lang, Ruth Duerr, Sarah Elizabeth West, Arika Virapongse, Christopher Kyba, Anne Bowser, Caren Cooper, Andrea Sforzi, et al. Citizen science terminology matters: Exploring key terms. *Citizen Science: Theory and Practice*, pages 1–20, 2017.

Alexandra Eveleigh, Charlene Jennett, Ann Blandford, Philip Brohan, and Anna L Cox. Designing for dabblers and deterring drop-outs in citizen science. In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems*, pages 2985–2994. ACM, 2014.

Alexandra Eveleigh, Charlene Jennett, Stuart Lynn, and Anna L Cox. "i want to be a captain! i want to be a captain!": Gamification in the old weather citizen science project. In *Proceedings of the first international conference on gameful design, research, and applications*, pages 79–82. ACM, 2013.

Morten W Fagerland. t-tests, non-parametric tests, and large studies—a paradox of statistical practice? *BMC Medical Research Methodology*, 12(1):78, 2012.

Jose Luis Fernandez-Marquez, Ioannis Charalampidis, Oula Abu-Amsha, SP Mohanty, Rosy Mondardini, Daniel K Schneider, Francois Grey, and Ben Segal. Ccltracker framework: Monitoring user engage-ment and learning in web-based citizen science projects. *Human Computation*, (3):99–117, 2016.

Debra A Fischer, Megan E Schwamb, Kevin Schawinski, Chris Lintott, John Brewer, Matt Giguere, Stuart Lynn, Michael Parrish, Thibault Sartori, Robert Simpson, et al. Planet hunters: the first two planet candidates identified by the public using the kepler public archive data. *Monthly Notices of the Royal Astronomical Society*, 419(4):2900–2911, 2012.

Gregory A Fredricks and Roger B Nelsen. On the relationship between spearman's rho and kendall's tau for pairs of continuous random variables. *Journal of statistical planning and inference*, 137(7):2143–2150, 2007.

Amy Freitag and Max J Pfeffer. Process, not product: investigating recommendations for improving citizen science "success". *PLoS One*, 8(5):e64079, 2013.

Catherine O Fritz, Peter E Morris, and Jennifer J Richler. Effect size estimates: current use, calculations, and interpretation. *Journal of experimental psychology: General*, 141(1):2, 2012.

Melanie A Galloway, Kyle W Willett, Lucy F Fortson, Carolin N Cardamone, Kevin Schawinski, Edmond Cheung, Chris J Lintott, Karen L Masters, Thomas Melvin, and Brooke D Simmons. Galaxy zoo: the effect of bar-driven fuelling on the presence of an active galactic nucleus in disc galaxies. *Monthly Notices of the Royal Astronomical Society*, 448(4):3442–3454, 2015.

Mary M Gardiner, Leslie L Allee, Peter MJ Brown, John E Losey, Helen E Roy, and Rebecca Rice Smyth. Lessons from lady beetles: accuracy of monitoring data from us and uk citizen-science programs. *Frontiers in Ecology and the Environment*, 10(9):471–476, 2012.

Andrew R Gilpin. Table for conversion of kendall's tau to spearman's rho within the context of measures of magnitude of effect for meta-analysis. *Educational and psychological measurement*, 53(1):87–92, 1993.

Christina Goulding. Grounded theory, ethnography and phenomenology: A comparative analysis of three qualitative strategies for marketing research. *European journal of Marketing*, 39(3/4):294–308, 2005.

J Greene. Mixing methods in social enquiry, 2007.

Anita Greenhill, Kate Holmes, Chris Lintott, Brooke Simmons, Karen Masters, Joe Cox, and Gary Graham. Playing with science: gamised aspects of gamification found on the online citizen science project-zooniverse. In *15th International Conference on Intelligent Games and Simulation.*, GAME-ON, pages 15–24. European Multidisciplinary Society for Modelling and Simulation Technology EUROSIS, 2014. ISBN 978-90-77381-85-4.

Wolfgang Grichting. Do storks really bring babies? *Journal of Tertiary Educational Administration*, 14(1):75–86, 1992.

N Gugliucci, P Gay, and G Bracey. Citizen science motivations as discovered with cosmoquest. In *Ensuring Stem Literacy: A National Conference on STEM Education and Public Outreach*, volume 483, page 437, 2014.

Mordechai E Haklay. Why is participation inequality important? In *European Handbook of Crowdsourced Geographic Information*, chapter 3. Ubiquity Press, 2016.

Muki Haklay. Citizen science and volunteered geographic information: Overview and typology of participation. In *Crowdsourcing geographic knowledge*, pages 105–122. Springer, 2013.

Muki Haklay. Citizen science and policy: a European perspective. *The Wodrow Wilson Center, Commons Lab*, 2015.

Russell Hardin. Collective action as an agreeable n-prisoners' dilemma. *Behavioral science*, 16(5):472–481, 1971.

Casper Harteveld, Amy Stahl, Gillian Smith, Cigdem Talgar, and Steven C Sutherland. Standing on the shoulders of citizens: Exploring gameful collaboration for creating social experiments. In *System Sciences (HICSS), 2016 49th Hawaii International Conference on*, pages 74–83. IEEE, 2016.

Joshua K Hartshorne, Claire Bonial, and Martha Palmer. The verbcorner project: Toward an empirically-based semantic decomposition of verbs. In *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing*, pages 1438–1442, 2013.

Lorna Heaton and Ricardo Vidal Torres. When does leisure become work? *Akademisk Kvarter (Academic Quarter)*, 11:85–19, 2015.

Susanne Hecker, Muki Haklay, Anne Bowser, Zen Makuch, and Johannes Vogel. *Citizen science: innovation in open science, society and policy*. UCL Press, 2018.

James Hollan, Edwin Hutchins, and David Kirsh. Distributed cognition: toward a new foundation for human-computer interaction research. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 7(2):174–196, 2000.

MW Home. Diy zooniverse citizen science project: engaging the public with your museum's collections and data. 2019.

Jeff Howe. *Crowdsourcing: How the power of the crowd is driving the future of business*. Random House, 2008.

Kenneth R Howe. Against the quantitative-qualitative incompatibility thesis or dogmas die hard. *Educational Researcher*, pages 10–16, 1988.

Chang Hu, Benjamin B Bederson, and Philip Resnik. Translation by iterative collaboration between monolingual users. In *Proceedings of Graphics Interface 2010*, pages 39–46. Canadian Information Processing Society, 2010.

Ioanna Iacovides, Charlene Jennett, Cassandra Cornish-Trestrail, and Anna L Cox. Do games attract or sustain engagement in citizen science?: a study of volunteer motivations. In *CHI'13 Extended Abstracts on Human Factors in Computing Systems*, pages 1101–1106. ACM, 2013.

Alicia Iriberri and Gondy Leroy. A life-cycle perspective on online community success. *ACM Computing Surveys (CSUR)*, 41(2):11, 2009.

Alan Irwin. Citizen science and scientific citizenship: Same words, different meanings? *Science Communication Today–2015*, pages 29–38, 2015.

Corey Jackson and Kevin Crowston Gabriel Mugar Katie DeVries. Modeling future work in crowdsourced citizen science, 2015.

Corey Jackson, Gabriel Mugar, Kevin Crowston, and Carsten Østerlund. Encouraging work in citizen science: Experiments in goal setting and anchoring. In *Proceedings of the 19th ACM Conference on Computer Supported Cooperative Work and Social Computing Companion*, pages 297–300. ACM, 2016a.

Corey Jackson, Carsten Østerlund, Kevin Crowston, Mahboobeh Harandi, and Gravity Spy Team. Supporting crowd workers: Assembling resources in online citizen science projects. In *Companion of the 2018 ACM Conference on Computer Supported Cooperative Work and Social Computing (To Appear)*. ACM, 2017.

Corey Jackson, Carsten Østerlund, Kevin Crowston, Gabriel Mugar, and KD Hassman. Motivations for sustained participation in citizen science: Case studies on the role of talk. In *17th ACM Conference on Computer Supported Cooperative Work & Social Computing*, 2014.

Corey Jackson, Carsten Østerlund, Veronica Maidel, Kevin Crowston, and Gabriel Mugar. Which way did they go?: Newcomer movement through the zooniverse. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*, pages 624–635. ACM, 2016b.

Corey Brian Jackson, Kevin Crowston, Gabriel Mugar, and Carsten Østerlund. "guess what! you're the first to see this event": Increasing contribution to online production communities. In *Proceedings of the 19th International Conference on Supporting Group Work*, GROUP '16, pages 171–179, New York, NY, USA, 2016c. ACM. ISBN 978-1-4503-4276-6.

Corey Brian Jackson, Kevin Crowston, and Carsten Østerlund. Working in the shadows: Anonymous contributions by users in citizen science. In *21st ACM Conference on Computer Supported Cooperative Work & Social Computing*, 2018.

Shaili Jain and David C Parkes. The role of game theory in human computation systems. In *Proceedings of the ACM SIGKDD Workshop on Human Computation*, pages 58–61. ACM, 2009.

Carlos M Jarque and Anil K Bera. Efficient tests for normality, homoscedasticity and serial independence of regression residuals. *Economics letters*, 6(3):255–259, 1980.

Caroline Jay, Robert Dunne, David Gelsthorpe, and Markel Vigo. To sign up, or not to sign up?: Maximizing citizen science contribution rates through optional registration. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, pages 1827–1832. ACM, 2016.

CHARLENE Jennett, DOMINIC Furniss, IOANNA Iacovides, SARAH Wiseman, SANDY JJ Gould, and ANNA L Cox. Exploring citizen psych-science and the motivations of errordiary volunteers. *Human Computation*, 1(2):200–218, 2014.

Charlene Jennett, Laure Kloetzer, Margaret Gold, and Anna L Cox. Sociability in virtual citizen science. In *Proceedings of the CHI 2013 workshop on designing and evaluating sociability in online video games*, pages 29–32. Self-published, 2013.

Charlene Jennett, Laure Kloetzer, Daniel Schneider, Ioanna Iacovides, Anna Cox, Margaret Gold, Brian Fuchs, Alexandra Eveleigh, Kathleen Methieu, Zoya Ajani, et al. Motivations, learning and creativity in online citizen science. *Journal of Science Communication*, 15(3), 2016.

Yuan Jia, Yikun Liu, Xing Yu, and Stephen Voida. Designing leaderboards for gamification: Perceived differences based on user ranking, application domain, and personality traits. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pages 1949–1960. ACM, 2017.

Yuan Jia, Bin Xu, Yamini Karanam, and Stephen Voida. Personality-targeted gamification: a survey study on personality traits and motivational affordances. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, pages 2001–2013. ACM, 2016.

Norman J Johnson. Modified t tests and confidence intervals for asymmetrical populations. *Journal of the American Statistical Association*, 73(363):536–544, 1978.

R Burke Johnson and Anthony J Onwuegbuzie. Mixed methods research: A research paradigm whose time has come. *Educational researcher*, 33(7):14–26, 2004.

Sean F Johnston, Benjamin Franks, and Sandy Whitelaw. Crowd-sourced science. *journal of Information Ethics*, 2017.

Joseph M Juran. Universals in management planning and controlling. *Management Review*, 43(11):748–761, 1954.

Geoff Kaufman, Mary Flanagan, and Sukdith Punjasthitkul. Investigating the impact of'emphasis frames' and social loafing on player motivation and performance in a crowdsourcing game. In *Proceedings of the 2016 CHI conference on human factors in computing systems*, pages 4122–4128. ACM, 2016.

Firas Khatib, Seth Cooper, Michael D Tyka, Kefan Xu, Ilya Makedon, Zoran Popović, David Baker, and Foldit Players. Algorithm discovery by protein folding game players. *Proceedings of the National Academy of Sciences*, 108(47):18949–18953, 2011.

Jinseop S Kim, Matthew J Greene, Aleksandar Zlateski, Kisuk Lee, Mark Richardson, Srinivas C Turaga, Michael Purcaro, Matthew Balkam, Amy Robinson, Bardia F Behabadi, et al. Space-time wiring specificity supports direction selectivity in the retina. *Nature*, 509(7500):331, 2014.

Aya H Kimura and Abby Kinchy. Citizen science: Probing the virtues and contexts of participatory research. *Engaging Science, Technology, and Society*, 2: 331–361, 2016.

Barbara Kitchenham. Procedure for undertaking systematic reviews. *Computer Science Depart-ment, Keele University (TRISE-0401) and National ICT Australia Ltd (0400011T. 1), Joint Technical Report*, 2004.

Barbara Kitchenham, O Pearl Brereton, David Budgen, Mark Turner, John Bailey, and Stephen Linkman. Systematic literature reviews in software engineering–a systematic literature review. *Information and software technology*, 51(1):7–15, 2009.

Laure Kloetzer, Daniel Schneider, Charlene Jennett, Ioanna Iacovides, Alexandra Eveleigh, Anna Cox, and Margaret Gold. Learning by volunteer computing, thinking and gaming: What and how are volunteers learning by participating in virtual citizen science? *Changing Configurations of Adult Education in Transitional Times*, page 73, 2014.

Hiromi Kobori, Janis L Dickinson, Izumi Washitani, Ryo Sakurai, Tatsuya Amano, Naoya Komatsu, Wataru Kitamura, Shinichi Takagawa, Kazuo Koyama, Takao Ogawara, et al. Citizen science: a new approach to advance ecology, education, and conservation. *Ecological Research*, 31(1):1–19, 2016.

Margaret Kosmala, Andrea Wiggins, Alexandra Swanson, and Brooke Simmons. Assessing data quality in citizen science. *Frontiers in Ecology and the Environment*, 14(10):551–560, 2016.

Helena Chmura Kraemer. On estimation and hypothesis testing problems for correlation coefficients. *Psychometrika*, 40(4):473–485, 1975.

Robert E Kraut, Paul Resnick, Sara Kiesler, Moira Burke, Yan Chen, Niki Kittur, Joseph Konstan, Yuqing Ren, and John Riedl. *Building successful online communities: Evidence-based social design.* Mit Press, 2012.

Donna M Kridelbaugh. The use of online citizen-science projects to provide experiential learning opportunities for nonmajor science students. *Journal of microbiology & biology education*, 17(1):105, 2016.

Thomas S Kuhn. The structure of scientific revolutions, 2nd. *Chicago: Univ. of Chicago Pr*, 1970.

Christopher Kullenberg and Dick Kasperowski. What is citizen science?–a scientometric meta-analysis. *PloS one*, 11(1):e0147152, 2016.

Janaki Kumar. Gamification at work: Designing engaging business software. In *International Conference of Design, User Experience, and Usability*, pages 528–537. Springer, 2013.

Daniël Lakens. Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and anovas. *Frontiers in psychology*, 4:863, 2013.

Cliff Lampe and Erik Johnston. Follow the (slash) dot: effects of feedback on new members in an online community. In *Proceedings of the 2005 international ACM SIGGROUP conference on Supporting group work*, pages 11–20. ACM, 2005.

John Landers. Quantification in history, topic 4: Hypothesis testing ii-differing central tendency, 1981.

Mirella Lapata. Automatic evaluation of information ordering: Kendall's tau. *Computational Linguistics*, 32(4):471–484, 2006.

Bruno Latour. Nonhumans. *Patterned ground: Entanglements of nature and culture*, pages 224–227, 2004.

Jeffrey Laut, Francesco Cappa, Oded Nov, and Maurizio Porfiri. Increasing citizen science contribution using a virtual peer. *Journal of the Association for Information Science and Technology*, 2016.

Jean Lave and Etienne Wenger. Legitimate peripheral participation. *Learners, learning and assessment, London: The Open University*, pages 83–89, 1999.

John Law. Notes on the theory of the actor-network: Ordering, strategy, and heterogeneity. *Systems practice*, 5(4):379–393, 1992.

Jeehyung Lee, Wipapat Kladwang, Minjae Lee, Daniel Cantu, Martin Azizyan, Hanjoo Kim, Alex Limpaecher, Sungroh Yoon, Adrien Treuille, and Rhiju Das. Rna design rules from a massive open laboratory. *Proceedings of the National Academy of Sciences*, 111(6):2122–2127, 2014.

Tae Kyoung Lee, Kevin Crowston, Mahboobeh Harandi, Carsten Østerlund, and Grant Miller. Appealing to different motivations in a message to recruit citizen scientists: results of a field experiment. *JCOM*, 17(01):A02–2, 2018.

Tae Kyoung Lee, Kevin Crowston, Carsten Østerlund, and Grant Miller. Recruiting messages matter: Message strategies to attract citizen scientists. In *Companion of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*, pages 227–230. ACM, 2017.

Andreas Lieberoth, Mads Kock Pedersen, Andreea Catalina Marin, Tilo Planke, and Jacob Friis Sherson. Getting humans to do quantum optimization-user acquisition, engagement and early results from the citizen cyberscience game quantum moves. *Human Computation*, 1(2), 2014.

Chris Lintott and Jason Reed. Human computation in citizen science. In *Handbook of human computation*, pages 153–162. Springer, 2013.

Chris J Lintott, Kevin Schawinski, William Keel, Hanny Van Arkel, Nicola Bennert, Edward Edmondson, Daniel Thomas, Daniel JB Smith, Peter D Herbert, Matt J Jarvis, et al. Galaxy zoo:'hanny's voorwerp', a quasar light echo? *Monthly Notices of the Royal Astronomical Society*, 399(1):129–140, 2009.

Chris J Lintott, Kevin Schawinski, Anže Slosar, Kate Land, Steven Bamford, Daniel Thomas, M Jordan Raddick, Robert C Nichol, Alex Szalay, Dan Andreescu, et al. Galaxy zoo: morphologies derived from visual inspection of galaxies from the sloan digital sky survey. *Monthly Notices of the Royal Astronomical Society*, 389(3):1179–1189, 2008.

Greg Little, Lydia B Chilton, Max Goldman, and Robert C Miller. Exploring iterative and parallel human computation processes. In *Proceedings of the ACM SIGKDD workshop on human computation*, pages 68–76. ACM, 2010.

Jeffrey D Long and Norman Cliff. Confidence intervals for kendall's tau. *British Journal of Mathematical and Statistical Psychology*, 50(1):31–41, 1997.

Markus Luczak-Roesch, Ramine Tinati, Elena Simperl, Max Van Kleek, Nigel Shadbolt, and Robert Simpson. Why won't aliens talk to us? content and community dynamics in online citizen science. In *Proceedings of the Eighth International AAAI Conference on Weblogs and Social Media*, 2014.

R Lukyanenko, J Parsons, and Y F Wiersma. The impact of task design on accuracy, completeness and discovery in surveillance-based crowdsourcing. In *Proceedings of the 4th International Conference on Collective Intelligence*, 2016.

Sylvia Luu and Anupama Narayan. Games at work: Examining a model of team effectiveness in an interdependent gaming task. *Computers in Human Behavior*, 2017.

Thomas W MacFarland and Jan M Yates. Mann–whitney u test. In *Introduction to nonparametric statistics for the biological sciences using R*, pages 103–132. Springer, 2016.

Michael J Madison. Commons at the intersection of peer production, citizen science, and big data: Galaxy zoo. *Oxford University Press*, 2014.

Henry B Mann and Donald R Whitney. On a test of whether one of two random variables is stochastically larger than the other. *The annals of mathematical statistics*, pages 50–60, 1947.

Andrew Mao, Ece Kamar, Yiling Chen, Eric Horvitz, Megan E Schwamb, Chris J Lintott, and Arfon M Smith. Volunteering versus work for pay: Incentives and tradeoffs in crowdsourcing. In *First AAAI conference on human computation and crowdsourcing*, 2013a.

Andrew Mao, Ece Kamar, and Eric Horvitz. Why stop now? predicting worker engagement in online crowdsourcing. In *First AAAI Conference on Human Computation and Crowdsourcing*, 2013b.

Gerald Marwell and Pamela Oliver. *The critical mass in collective action*. Cambridge University Press, 1993.

Gerald Marwell, Pamela E Oliver, and Ralph Prahl. Social networks and collective action: A theory of the critical mass. iii. *American Journal of Sociology*, 94 (3):502–534, 1988.

Aaron D Mason, Georgios Michalakidis, and Paul J Krause. Tiger nation: Empowering citizen scientists. In *Digital Ecosystems Technologies (DEST), 2012 6th IEEE International Conference on*, pages 1–5. IEEE, 2012.

K Masters, EY Oh, J Cox, B Simmons, C Lintott, G Graham, A Greenhill, and K Holmes. Science learning via participation in online citizen science. *Journal of Science Communication*, 15(3):A07–A07, 2016.

Fiona McElduff, Mario Cortina-Borja, Shun-Kai Chan, and Angie Wade. When t-tests or wilcoxon-mann-whitney tests won't do. *Advances in physiology education*, 34(3):128–133, 2010.

Elisa D Mekler, Florian Brühlmann, Klaus Opwis, and Alexandre N Tuch. Disassembling gamification: the effects of points and meaning on user motivation and performance. In *CHI'13 extended abstracts on human factors in computing systems*, pages 1137–1142. ACM, 2013a.

Elisa D Mekler, Florian Brühlmann, Klaus Opwis, and Alexandre N Tuch. Do points, levels and leaderboards harm intrinsic motivation?: an empirical analysis of common gamification elements. In *Proceedings of the First International Conference on gameful design, research, and applications*, pages 66–73. ACM, 2013b.

Pietro Michelucci. *Handbook of Human Computation*. Springer Science & Business Media, 2013.

Abraham Miller-Rushing, Richard Primack, and Rick Bonney. The history of public participation in ecological research. *Frontiers in Ecology and the Environment*, 10(6):285–290, 2012.

A.M.M. Morais, J. Raddick, and R.D.C. Santos. Visualization and characterization of users in a citizen science project. In BD Broome, DL Hall, and J Llinas, editors, *NEXT-GENERATION ANALYST*, volume 8758 of *Proceedings of SPIE*. SPIE, 2013. ISBN 978-0-8194-9549-5. Conference on Next-Generation Analyst, Baltimore, MD, APR 29-30, 2013.

A.M.M. Morais, R.D.C. Santos, and M.J. Raddick. Visualization of citizen science volunteers' behaviors with data from usage logs. *Computing in Science and Engineering*, 17(4):42–50, 2015.

David L Morgan. Paradigms lost and pragmatism regained: Methodological implications of combining qualitative and quantitative methods. *Journal of mixed methods research*, 1(1):48–76, 2007.

Gabriel Mugar, Carsten Østerlund, Katie DeVries Hassman, Kevin Crowston, and Corey Brian Jackson. Planet hunters and seafloor explorers: legitimate peripheral participation through practice proxies in online citizen science. In *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing*, pages 109–119. ACM, 2014.

Gabriel Mugar, Carsten Østerlund, Corey Brian Jackson, and Kevin Crowston. Being present in online communities: learning in citizen science. In *Proceedings of the 7th International Conference on Communities and Technologies*, pages 129–138. ACM, 2015.

Rami Wael Muhtaseb. Digital collaboration in educational and research institutions. In *Crowdsourcing: Concepts, Methodologies, Tools, and Applications*, pages 649–662. IGI Global, 2019.

Mavuto M Mukaka. A guide to appropriate use of correlation coefficient in medical research. *Malawi Medical Journal*, 24(3):69–71, 2012.

Jacqueline Murray. Likert data: what to use, parametric or non-parametric? *International Journal of Business and Social Science*, 4(11), 2013.

Nadim Nachar et al. The mann-whitney u: A test for assessing whether two independent samples come from the same distribution. *Tutorials in Quantitative Methods for Psychology*, 4(1):13–20, 2008.

Bonnie A Nardi. Coda and response to christine halverson. *Computer Supported Cooperative Work (CSCW)*, 11(1-2):269–275, 2002.

Greg Newman, Andrea Wiggins, Alycia Crall, Eric Graham, Sarah Newman, and Kevin Crowston. The future of citizen science: emerging technologies and shifting paradigms. *Frontiers in Ecology and the Environment*, 10(6):298–304, 2012.

Geoff Norman. Likert scales, levels of measurement and the "laws" of statistics. *Advances in health sciences education*, 15(5):625–632, 2010.

Oded Nov, Ofer Arazy, and David Anderson. Dusting for science: motivation and participation of digital citizen science volunteers. In *Proceedings of the 2011 iConference*, pages 68–74. ACM, 2011a.

Oded Nov, Ofer Arazy, and David Anderson. Technology-mediated citizen science participation: A motivational model. In *ICWSM*, 2011b.

Oded Nov, Ofer Arazy, and David Anderson. Scientists home: what drives the quantity and quality of online citizen science participation. *PloS one*, 9(4):e90375, 2014.

Ryan P O'Donnell and Andrew M Durso. Harnessing the power of a global network of citizen herpetologists by improving citizen science databases. *Herpetological Review*, 45(1):151–157, 2014.

Pamela Oliver. " if you don't do it, nobody else will": Active and token contributors to local collective action. *American sociological review*, pages 601–610, 1984.

Mancur Olson. *Logic of collective action: Public goods and the theory of groups (Harvard economic studies. v. 124)*. Harvard University Press, 1965.

Elinor Ostrom. A behavioral approach to the rational choice theory of collective action: Presidential address, american political science association, 1997. *American political science review*, 92(1):1–22, 1998.

Yue Pan and Eli Blevis. A survey of crowdsourcing as a means of collaboration and the implications of crowdsourcing for interaction design. In *Collaboration Technologies and Systems (CTS), 2011 International Conference on*, pages 397–403. IEEE, 2011.

Rajul E Pandya. A framework for engaging diverse communities in citizen science in the us. *Frontiers in Ecology and the Environment*, 10(6):314–317, 2012.

Charles S Peirce. What pragmatism is. *The Monist*, 15(2):161–181, 1905.

Lesandro Ponciano and Francisco Brasileiro. Finding volunteers' engagement profiles in human computation for citizen science projects. *arXiv preprint arXiv:1501.02134*, 2015.

Lesandro Ponciano, Francisco Brasileiro, Robert Simpson, and Arfon Smith. Volunteers' engagement in human computation for astronomy projects. *Computing in Science & Engineering*, 16(6):52–59, 2014.

Marisa Ponti, Thomas Hillman, Christopher Kullenberg, and Dick Kasperowski. Getting it right or being top rank: Games in citizen science. *Citizen Science: Theory and Practice*, 3(1), 2018.

Jenny Preece. Sociability and usability in online communities: Determining and measuring success. *Behaviour & Information Technology*, 20(5):347–356, 2001.

Nathan Prestopnik, Kevin Crowston, and Jun Wang. Exploring data quality in games with a purpose. *iConference 2014 Proceedings*, 2014.

Nathan Prestopnik, Kevin Crowston, and Jun Wang. Gamers, citizen scientists, and data: Exploring participant contributions in two games with a purpose. *Computers in Human Behavior*, 68:254–268, 2017.

Nathan R Prestopnik and Kevin Crowston. Citizen science system assemblages: understanding the technologies that support crowdsourced science. In *Proceedings of the 2012 iConference*, pages 168–176. ACM, 2012.

Nathan R Prestopnik and Jian Tang. Points, stories, worlds, and diegesis: Comparing player experiences in two citizen science games. *Computers in Human Behavior*, 52:492–506, 2015.

Steven J Price and Michael E Dorcas. The carolina herp atlas: an online, citizen-science approach to document amphibian and reptile occurrences. *Herpetological Conservation and Biology*, 6(2):287–296, 2011.

Alexander J Quinn and Benjamin B Bederson. A taxonomy of distributed human computation. *Human-Computer Interaction Lab Tech Report, University of Maryland*, 2009.

Alexander J Quinn and Benjamin B Bederson. Human computation: a survey and taxonomy of a growing field. In *Proceedings of the SIGCHI conference on human factors in computing systems*, pages 1403–1412. ACM, 2011.

Jordan Raddick, GL Bracey, and PL Gay. Motivations of citizen scientists participating in galaxy zoo: A more detailed look. In *Bulletin of the American Astronomical Society*, volume 42, page 509, 2010.

M Jordan Raddick, Georgia Bracey, Karen Carney, Geza Gyuk, Kirk Borne, John Wallin, Suzanne Jacoby, and Adler Planetarium. Citizen science: status and research directions for the coming decade. *AGB Stars and Related Phenomenastro 2010: The Astronomy and Astrophysics Decadal Survey*, page 46P, 2009a.

M Jordan Raddick, Georgia Bracey, Pamela L Gay, Chris J Lintott, Carie Cardamone, Phil Murray, Kevin Schawinski, Alexander S Szalay, and Jan Vandenberg. Galaxy zoo: Motivations of citizen scientists. *arXiv preprint arXiv:1303.6886*, 12(1), 2013.

M Jordan Raddick, Georgia Bracey, Pamela L Gay, Chris J Lintott, Phil Murray, Kevin Schawinski, Alexander S Szalay, and Jan Vandenberg. Galaxy zoo: Exploring the motivations of citizen science volunteers. *arXiv preprint arXiv:0909.2925*, 2009b.

Marilyn A Ray. A philosophical method to study nursing phenomena. *Qualitative research methods in nursing*, pages 81–92, 1985.

Jeff Reed, M Jordan Raddick, Andrea Lardner, and Karen Carney. An exploratory factor analysis of motivations for participating in zooniverse, a collection of virtual citizen science projects. In *System Sciences (HICSS), 2013 46th Hawaii International Conference on*, pages 610–619. IEEE, 2013.

DG Rees. Non-parametric hypothesis tests. In *Essential Statistics*, pages 116–129. Springer, 1989.

Neal Reeves. Sociality in virtual citizen science (presentation and panel discussion). Presented at 20th ACM Conference on Computer-Supported Cooperative Work and Social Computing in Portland, OR, USA on 25 February 2017, 2017.

Neal Reeves, Ramine Tinati, Sergej Zerr, Elena Simperl, and Max Van Kleek. 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*, pages 2137–2152. ACM, ACM, 2017.

Neal Reeves, Max Van Kleek, and Elena Simperl. From crowd to community: Support for community features in online citizen science projects. Presented at GESIS Computational Social Science Winter Symposium 2015 in Cologne, Germany on 2 December 2015, 2015.

Neal Reeves, Peter West, and Elena Simperl. ‘`a game without competition is hardly a game”: The impact of competitions on player activity in a human computation game. In *Proceedings of the AAAI Conference on Human Computation (HCOMP 2018)., HCOMP*, volume 18, 2018.

David Reisman. *Theories of collective action: Downs, Olson and Hirsch*. Springer, 1990.

Yuqing Ren, Robert Kraut, and Sara Kiesler. Applying common identity and bond theory to design of online communities. *Organization studies*, 28(3):377–408, 2007.

Marnie E Rice and Grant T Harris. Comparing effect sizes in follow-up studies: Roc area, cohen’s d, and r. *Law and human behavior*, 29(5):615–620, 2005.

Catherine Ridings and Molly McLure Wasko. Online discussion group sustainability: Investigating the interplay between structural dynamics and social dynamics over time. *Journal of the Association for Information Systems*, 11(2): 1, 2010.

Catherine M Ridings and David Gefen. Virtual community attraction: Why people hang out online. *Journal of Computer-Mediated Communication*, 10(1): 00–00, 2004.

Stephen Robertson, Milan Vojnovic, and Ingmar Weber. Rethinking the esp game. In *CHI’09 Extended Abstracts on Human Factors in Computing Systems*, pages 3937–3942. ACM, 2009.

Yvonne Rogers. Hci theory: classical, modern, and contemporary. *Synthesis Lectures on Human-Centered Informatics*, 5(2):1–129, 2012.

Yvonne Rogers and Judi Ellis. Distributed cognition: an alternative framework for analysing and explaining collaborative working. *Journal of information technology*, 9(2):119–128, 1994.

Richard Rorty. *Consequences of pragmatism: Essays, 1972-1980*. U of Minnesota Press, 1982.

Murray Rosenblatt. A central limit theorem and a strong mixing condition. *Proceedings of the National Academy of Sciences*, 42(1):43–47, 1956.

Holly Rosser and Andrea Wiggins. Crowds and camera traps: Genres in online citizen science projects. In *Proceedings of the 52nd Hawaii International Conference on System Sciences*, 2019.

Dana Rotman, Jen Hammock, Jenny Preece, Derek Hansen, Carol Boston, Anne Bowser, and Yurong He. Motivations affecting initial and long-term participation in citizen science projects in three countries. *iConference 2014 Proceedings*, 2014.

Dana Rotman, Jenny Preece, Jen Hammock, Kezee Procita, Derek Hansen, Cynthia Parr, Darcy Lewis, and David Jacobs. Dynamic changes in motivation in collaborative citizen-science projects. In *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work*, pages 217–226. ACM, 2012.

Jeffrey N Rouder, Paul L Speckman, Dongchu Sun, Richard D Morey, and Geoffrey Iverson. Bayesian t tests for accepting and rejecting the null hypothesis. *Psychonomic bulletin & review*, 16(2):225–237, 2009.

HE Roy, MJO Pocock, CD Preston, DB Roy, J Savage, JC Tweddle, and LD Robinson. Understanding citizen science and environmental monitoring: final report on behalf of UK Environmental Observation Framework. NERC/Centre for Ecology & Hydrology, 2012.

Melvin T Rupinski and William P Dunlap. Approximating pearson product-moment correlations from kendall's tau and spearman's rho. *Educational and psychological measurement*, 56(3):419–429, 1996.

Richard M Ryan and Edward L Deci. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American psychologist*, 55(1):68, 2000.

Michael Sailer, Jan Ulrich Hense, Sarah Katharina Mayr, and Heinz Mandl. How gamification motivates: An experimental study of the effects of specific game design elements on psychological need satisfaction. *Computers in Human Behavior*, 69:371–380, 2017.

Yasuaki Sakamoto, Yuko Tanaka, Lixiu Yu, and Jeffrey V Nickerson. The crowdsourcing design space. In *International Conference on Foundations of Augmented Cognition*, pages 346–355. Springer, 2011.

Henry Sauermann and Chiara Franzoni. Crowd science user contribution patterns and their implications. *Proceedings of the National Academy of Sciences*, 112(3): 679–684, 2015.

Shlomo S Sawilowsky. New effect size rules of thumb. *Journal of Modern Applied Statistical Methods*, 8(2):26, 2009.

Edwin Sayes. Actor–network theory and methodology: Just what does it mean to say that nonhumans have agency? *Social Studies of Science*, 44(1):134–149, 2014.

Katie Seaborn and Deborah I Fels. Gamification in theory and action: A survey. *International Journal of Human-Computer Studies*, 74:14–31, 2015.

Avi Segal, Ya'akov Kobi Gal, Robert J Simpson, Victoria Victoria Homsy, Mark Hartswood, Kevin R Page, and Marina Jirotka. Improving productivity in citizen science through controlled intervention. In *Proceedings of the 24th International Conference on World Wide Web Companion*, pages 331–337. International World Wide Web Conferences Steering Committee, 2015.

Nigel R Shadbolt, Daniel A Smith, Elena Simperl, Max Van Kleek, Yang Yang, and Wendy Hall. Towards a classification framework for social machines. In *Proceedings of the 22nd International Conference on World Wide Web*, pages 905–912. ACM, 2013.

Sally Shuttleworth. Life in the zooniverse: working with citizen science. *Journal of Literature and Science*, 10(1):46–51, 2017.

Jonathan Silvertown. A new dawn for citizen science. *Trends in ecology & evolution*, 24(9):467–471, 2009.

Jonathan Silvertown, Martin Harvey, Richard Greenwood, Mike Dodd, Jon Rosewell, Tony Rebelo, Janice Ansine, and Kevin McConway. Crowdsourcing the identification of organisms: A case-study of ispot. *ZooKeys*, 480:125, 2015.

Elena Simperl, Neal Reeves, Christopher Phethean, Todd Lynes, and Ramine Tinati. Is virtual citizen science a game? *Transactions on Social Computing*, 2018.

Kristin Siu, Alexander Zook, and Mark O Riedl. Collaboration versus competition: Design and evaluation of mechanics for games with a purpose. In *FDG*, 2014.

Paul Smart. Situating machine intelligence within the cognitive ecology of the internet. *Minds and Machines*, 27(2):357–380, 2017.

Paul Smart, Elena Simperl, and Nigel Shadbolt. A taxonomic framework for social machines. In *Social Collective Intelligence*, pages 51–85. Springer, 2014.

Paul R Smart and Nigel R Shadbolt. Social machines. In *Encyclopedia of Information Science and Technology*, chapter 675, pages 6855–6862. IGI Global, third edition, 2014.

Kevin Sparks, Alexander Klippel, Jan Oliver Wallgrün, and David Mark. Citizen science land cover classification based on ground and aerial imagery. In *International Workshop on Spatial Information Theory*, pages 289–305. Springer, 2015.

Helen Spiers, Alexandra Swanson, Lucy Fortson, Brooke Simmons, Laura Trouille, Samantha Blickhan, and Chris Lintott. Everyone counts? design considerations in online citizen science. *Journal of Science Communication*, 18(1), 2019.

Helen Spiers, Alexandra Swanson, Lucy Fortson, Brooke D Simmons, Laura Trouille, Samantha Blickhan, and Chris Lintott. Patterns of volunteer behaviour across online citizen science. In *Companion of the The Web Conference 2018 on The Web Conference 2018*, pages 93–94. International World Wide Web Conferences Steering Committee, 2018.

James Sprinks, Jeremy Morley, Steven Bamford, and Robert Houghton. The impact of task workflow design on vgi citizen science platforms, 2014.

James Sprinks, Jessica Wardlaw, Robert Houghton, Steven Bamford, and Jeremy Morley. Task workflow design and its impact on performance and volunteers' subjective preference in virtual citizen science. *International Journal of Human-Computer Studies*, 104:50–63, 2017.

Jared Starr, Charles M Schweik, Nathan Bush, Lena Fletcher, Jack Finn, Jennifer Fish, and Charles T Bargeron. Lights, camera… citizen science: assessing the effectiveness of smartphone-based video training in invasive plant identification. *PloS one*, 9(11):e111433, 2014.

Alexandra Swanson, Margaret Kosmala, Chris Lintott, Robert Simpson, Arfon Smith, and Craig Packer. Snapshot serengeti, high-frequency annotated camera trap images of 40 mammalian species in an african savanna. *Scientific data*, 2, 2015.

Abbas Tashakkori and Charles Teddlie. *Sage handbook of mixed methods in social & behavioral research*. Sage, 2010.

Arthur Tatnall. Actor-network theory in information systems research. In *Encyclopedia of Information Science and Technology, First Edition*, pages 42–46. IGI Global, 2005.

Richard Taylor. Interpretation of the correlation coefficient: a basic review. *Journal of diagnostic medical sonography*, 6(1):35–39, 1990.

Charles Teddlie and Abbas Tashakkori. *Foundations of mixed methods research: Integrating quantitative and qualitative approaches in the social and behavioral sciences*. Sage, 2009.

Omer Tene and Jules Polonetsky. Big data for all: Privacy and user control in the age of analytics. *Nw. J. Tech. & Intell. Prop.*, 11:xxvii, 2012.

EJ Theobald, AK Ettinger, HK Burgess, LB DeBey, NR Schmidt, HE Froehlich, C Wagner, J HilleRisLambers, J Tewksbury, MA Harsch, et al. Global change and local solutions: Tapping the unrealized potential of citizen science for biodiversity research. *Biological Conservation*, 181:236–244, 2015.

Ramine Tinati, Markus Luczak-Roesch, Elena Simperl, and Wendy Hall. An investigation of player motivations in eyewire, a gamified citizen science project. *Computers in Human Behavior*, 73:527–540, 2017a.

Ramine Tinati, Markus Luczak-Rösch, E Simperl, Wendy Hall, and Nigel Shadbolt. /command'and conquer: analysing discussion in a citizen science game. In *ACM Web Science Conference 2015*, page 26. ACM, 2015a.

Ramine Tinati, Elena Simperl, and Markus Luczak-Roesch. To help or hinder: Real-time chat in citizen science. *Proceedings of ICWSM*, 2017, 2017b.

Ramine Tinati, Elena Simperl, Markus Luczak-Roesch, Max Van Kleek, and Nigel Shadbolt. Collective intelligence in citizen science–a study of performers and talkers. *Proceedings of Collective Intelligence 2014*, 2014.

Ramine Tinati, Max Van Kleek, Elena Simperl, Markus Luczak-Rösch, Robert Simpson, and Nigel Shadbolt. Designing for citizen data analysis: A cross-sectional case study of a multi-domain citizen science platform. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, pages 4069–4078. ACM, 2015b.

Katrine Tovgaard-Olsen, Barbara Revuelta-Eugercios, and Allan Vestergaard. Providing open and public data: the challenges of user rights and the gdpr within the crowdsourcing projects at the danish national archives. 2019.

David Trafimow and Stephen Rice. A test of the null hypothesis significance testing procedure correlation argument. *The Journal of general psychology*, 136 (3):261–270, 2009.

Laura Trouille. From clicks to publications: How the public is changing the way we do research. In *Frank N. Bash Symposium 2015*, volume 261, page 009. SISSA Medialab, 2016.

Martijn Van Zomeren, Russell Spears, Agneta H Fischer, and Colin Wayne Leach. Put your money where your mouth is! explaining collective action tendencies through group-based anger and group efficacy. *Journal of personality and social psychology*, 87(5):649, 2004.

Jeff Vass and Jo E Munson. Revisiting the three rs of social machines: Reflexivity, recognition and responsivity. In *Proceedings of the 24th International Conference on World Wide Web*, pages 1161–1166. ACM, 2015.

Ben Vershbow. Nypl labs: Hacking the library. *Journal of Library Administration*, 53(1):79–96, 2013.

Benjamin Vershbow. Scribe: Turning text into structured information through the power of the crowd, 2015.

Luis Von Ahn and Laura Dabbish. Labeling images with a computer game. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 319–326. ACM, 2004.

Dara M Wald, Justin Longo, and AR Dobell. Design principles for engaging and retaining virtual citizen scientists. *Conservation Biology*, 2015.

David A Walker. Jmasm9: converting kendall's tau for correlational or meta-analytic analyses. *Journal of Modern Applied Statistical Methods*, 2(2):26, 2003.

James R Wallace, Saba Oji, and Craig Anslow. Technologies, methods, and values: Changes in empirical research at cscw 1990-2015. *Proceedings of the ACM on Human-Computer Interaction*, 1(CSCW):106, 2017.

Geoff Walsham. Actor-network theory and is research: current status and future prospects. In *Information systems and qualitative research*, pages 466–480. Springer, 1997.

David Watson and Luciano Floridi. Crowdsourced science: sociotechnical epistemology in the e-research paradigm. *Synthese*, 195(2):741–764, 2018.

Jane Webster and Richard T Watson. Analyzing the past to prepare for the future: Writing a literature review. *MIS quarterly*, pages xiii–xxiii, 2002.

Christopher C Werry. Internet relay chat. *Computer-mediated communication: Linguistic, social and cross-cultural perspectives*, pages 47–63, 1996.

Andrea Whittle and André Spicer. Is actor network theory critique? *Organization studies*, 29(4):611–629, 2008.

Andrea Wiggins and Kevin Crowston. From conservation to crowdsourcing: A typology of citizen science. In *System Sciences (HICSS), 2011 44th Hawaii international conference on*, pages 1–10. IEEE, 2011.

Andrea Wiggins and Kevin Crowston. Goals and tasks: Two typologies of citizen science projects. In *System Science (HICSS), 2012 45th Hawaii International Conference on*, pages 3426–3435. IEEE, 2012.

Andrea Wiggins and Yurong He. Community-based data validation practices in citizen science. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*, pages 1548–1559. ACM, 2016.

Andrea Wiggins, Greg Newman, Robert D Stevenson, and Kevin Crowston. Mechanisms for data quality and validation in citizen science. In *e-Science Workshops (eScienceW), 2011 IEEE Seventh International Conference on*, pages 14–19. IEEE, 2011.

Kyle W Willett, Chris J Lintott, Steven P Bamford, Karen L Masters, Brooke D Simmons, Kevin RV Casteels, Edward M Edmondson, Lucy F Fortson, Sugata Kaviraj, William C Keel, et al. Galaxy Zoo 2: detailed morphological classifications for 304 122 galaxies from the Sloan Digital Sky Survey. *Monthly Notices of the Royal Astronomical Society*, 435(4):2835–2860, 2013.

Danuta M Wojnar and Kristen M Swanson. Phenomenology: an exploration. *Journal of holistic nursing*, 25(3):172–180, 2007.

Chris Wood, Brian Sullivan, Marshall Iliff, Daniel Fink, and Steve Kelling. ebird: engaging birders in science and conservation. *PLoS Biol*, 9(12):e1001220, 2011.

Jamie Woodcock, Anita Greenhill, Kate Holmes, Gary Graham, Joe Cox, EY Oh, and Karen Masters. Crowdsourcing citizen science: Exploring the tensions between paid professionals and users. *Journal of Peer Production*, 2017.

James Wynn. *Citizen science in the digital age: rhetoric, science, and public engagement*. University of Alabama Press, 2017.

Weichao Xu, Yunhe Hou, YS Hung, and Yuexian Zou. A comparative analysis of spearman's rho and kendall's tau in normal and contaminated normal models. *Signal Processing*, 93(1):261–276, 2013.

Stephen C Yanchar and David D Williams. Reconsidering the compatibility thesis and eclecticism: Five proposed guidelines for method use. *Educational researcher*, 35(9):3–12, 2006.

Jun Yu, Weng-Keen Wong, Rebecca Hutchinson, et al. Modeling experts and novices in citizen science data for species distribution modeling. In *Data Mining (ICDM), 2010 IEEE 10th International Conference on*, pages 1157–1162. IEEE, 2010.

Jerrold H Zar. Biostatistical analysis. *Biostatistical Analysis (4th Edition)*, 55: 75309, 1996.

Haichao Zheng, Dahui Li, and Wenhua Hou. Task design, motivation, and participation in crowdsourcing contests. *International Journal of Electronic Commerce*, 15(4):57–88, 2011.

Xinxue Zhou, Jian Tang, Tianmei Wang, and Yanlin Ma. Investigating the impacts of task characteristics in gamified citizen science. *PACIS 2017 Proceedings*, 2017.

Appendix A

Round One Literature Review Results

A.1 Applications

These papers concerned only applications of VCS projects (e.g., for education)

Douglas Allchin. Using a free online citizen-science project to teach observation & quantification of animal behavior. *The american biology Teacher*, 72(9): 590–592, 2010.

Sarah K Anderson. *Bringing School to Life: Place-based Education Across the Curriculum.* Rowman & Littlefield, 2017.

Pauline Barmby, Jan Cami, and Sarah C Gallagher. " citizen science" projects for non-science astronomy students. 2011.

Daniel F Barringer, Julia D Plummer, Julia Kregenow, and Christopher Palma. Gamified approach to teaching introductory astronomy online. *Physical Review Physics Education Research*, 14 (1):10140, 2018.

Brigid Barron, Caitlin K Martin, Véronique Mertl, and Mohamed Yassine. Citizen science: Connecting to nature through networks. In *Mass collaboration and education*, pages 257–284. Springer, 2016.

Aerin Workman Benavides, Samuel Miller, and Catherine Matthews. *Meanings teachers make of teaching science outdoors as they explore citizen science.* University of North Carolina at Greensboro, 2016.

Isabelle Bonhoure, Toni Pou, Salvador Ferré, Núria Ferran-Ferrer, and Josep Perelló. High motivation and relevant scientic competencies through the introduction of citizen science at secondary schools: An assessment using a rubric model. In *Citizen Inquiry*, pages 168–193. Routledge, 2017.

Rick Bonney, Tina B Phillips, Heidi L Ballard, and Jody W Enck. Can citizen science enhance public understanding of science? *Public Understanding of Science*, 25(1):2–16, 2016.

Rick Bonney, Tina B Phillips, Jody Enck, Jennifer Shirk, and Nancy Trautmann. Citizen science and youth education. *National Research Council Committee on Out-of-School Time STEM. Washington, DC: National Research Council*, 2014.

KA Borden, A Kapadia, A Smith, and L Whyte. Educational exploration of the zooniverse: Tools for formal and informal audience engagement. In J Barnes, C Shupla, JG Manning, and MG Gibbs, editors, *Communicating Science: A National Conference on Science Education and Public Outreach*, volume 473 of *Astronomical Society of the Pacific Conference Series*, page 101. Astronom Soc Pacific, 2013. ISBN 978-1-58381-830-5. Conference on Communicating Science - A National Conference on Science Education and Public Outreach, Tucson, AZ, AUG 04-08, 2012.

Georgia Bracey. Teaching with citizen science: An exploratory study of teachers' motivations & perceptions. 2018.

Sanlyn Buxner, Martin Formanek, Chris David Impey, and Matthew Wenger. An analysis of learners in introductory astronomy massive open online courses. In *American Astronomical Society Meeting Abstracts# 228*, volume 228, 2016.

C. Cooper and A. Balakrishna. *Citizen Science Perspectives on E-Participation in Urban Planning.* 2013.

Caren Beth Cooper, Kerrie Anne Therese Loyd, Tessa Murante, Matthew Savoca, and Janis Dickinson. Natural history traits associated with detecting mortality within residential bird communities: can citizen science provide insights? *Environmental management*, 50 (1):11–20, 2012.

Kathy Costello, Ellen Reilly, Georgia Bracey, and Pamela Gay. Hanny and the mystery of the voorwerp: Citizen science in the classroom. In JB Jensen, JG Manning, MG Gibbs, and D Daon, editors, *CONNECTING PEOPLE TO SCIENCE: A NATIONAL CONFERENCE ON SCIENCE EDUCATION AND PUBLIC OUTREACH*, volume 457 of *Astronomical Society of the Pacific Conference Series*, pages 23–26. Space Telescope Sci Inst; NASA Lunar Sci Inst; NASAs Explorat Program; Infrared Process & Anal Ctr; NASAs Herschel Sci Ctr; Spitzer Sci Ctr; Stratospher Observ Infrared Astron; NASAs Chandra XRay Observ; Univ Chicago Press; Natl Radio Astron Observ; Ball Aerospace; Capitol Coll; Sky Skan; Univ Wyoming CAPER Team; Amer Astron Soc; AAS Educ Off; Solar Dynam Observ; Seiler Instrument; Explore Sci; Pratt St Ale House; MWT Assoc Inc; Amer Elements; Charlesbridge; Celestron, 2012. ISBN 978-1-58381-796-4. 123rd Annual Meeting of the Astronomical-Society-of-the-Pacific, Amer Geophys Union, Baltimore, MD, JUL 30-AUG 03, 2011.

A.L. Cushing. Using citizen science projects to develop cases for teaching digital curation. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 10766 LNCS:615–619, 2018.

P.D. da Silva and L. Heaton. Fostering digital and scientific literacy: Learning through practice. *First Monday*, 22(6), 2017.

D. Dailey and K. Starbird. "it's raining dispersants": Collective sensemaking of complex information in crisis contexts. volume 2015-January, pages 155–158, 2015.

Justin Dillon. Out-of-school science. *Encyclopedia of Science Education*, pages

727–728, 2015.

Ria Ann Dunkley. The role of citizen science in environmental education: A critical exploration of the environmental citizen science experience. In *Analyzing the Role of Citizen Science in Modern Research*, pages 213–230. IGI Global, 2017.

Lauren M Gardiner and Steven P Bachman. The role of citizen science in a global assessment of extinction risk in palms (arecaceae). *Botanical journal of the Linnean Society*, 182(2):543–550, 2016.

Pamela L Gay et al. Moving people from science adjacent to science doers with twitch. tv. In *AAS/Division for Planetary Sciences Meeting Abstracts*, volume 49, 2017.

M.J. Gaydos and K.D. Squire. Role playing games for scientific citizenship. *Cultural Studies of Science Education*, 7 (4):821–844, 2012.

MG Gibbs, S Buxner, P Gay, DA Crown, G Bracey, N Gugliucci, K Costello, and E Reilly. Collaboration in teacher workshops and citizen science. In *AGU Fall Meeting Abstracts*, 2013.

N Gugliucci, P Gay, G Bracey, K Costello, C Lehan, J Moore, and E Reilly. Citizen science and planetary science education through cosmoquest. In *AGU Fall Meeting Abstracts*, 2012.

ME Haklay. Beyond quantification: A role for citizen science and community science in a smart city. Data and City Workshop, 2015.

Hashima Hasan, Denise Smith, and Mangala Sharma. Nasa's astrophysics education and public outreach: Selected highlights. In J Barnes, C Shupla, JG Manning, and MG Gibbs, editors, *COMMUNICATING SCIENCE: A NATIONAL CONFERENCE ON SCIENCE EDUCATION AND PUBLIC OUTREACH*, volume 473 of *Astronomical Society of the Pacific Conference Series*, pages 145+. Astronom Soc Pacific, 2013. ISBN 978-1-58381-830-5. Conference on Communicating Science - A National Conference on Science Education and Public Outreach, Tucson, AZ, AUG 04-08, 2012.

S Henderson, D Ward, L Wasser, K Meymaris, and SJ Newman. Neon's citizen science academy: Exploring online professional development courses for educators to enhance participation. In *AGU Fall Meeting Abstracts*, 2012.

Christothea Herodotou. Citizen science and informal learning: A brief commentary. 2018.

Christothea Herodotou, Maria Aristeidou, Mike Sharples, and Eileen Scanlon. Designing citizen science tools for learning: lessons learnt from the iterative development of nquire. *Research and Practice in Technology Enhanced Learning*, 13(1):4, 2018.

Suzanne E Hiller and Anastasia Kitsantas. Fostering student metacognition and motivation in stem through citizen science programs. In *Metacognition: Fundaments, Applications, and Trends*, pages 193–221. Springer, 2015.

Z. Hossain, X. Jin, E.W. Bumbacher, A.M. Chung, S. Koo, J.D. Shapiro, C.Y. Truong, S. Choi, N.D. Orloff, P. Blikstein, and I.H. Riedel-Kruse. Interactive cloud experimentation for biology: An online education case study. In *CHI 2015: PROCEEDINGS OF THE 33RD ANNUAL CHI CONFERENCE ON HUMAN FACTORS IN COMPUTING SYSTEMS*, volume 2015-April, pages 3681–3690. Assoc Comp Machinery; Assoc Comp Machinery Special Interest Grp Comp Human Interact, 2015. ISBN 978-1-4503-3145-6. 33rd Annual CHI Conference on Human Factors in Computing Systems (CHI), Seoul, SOUTH KOREA, APR 18-23, 2015.

Zahid Hossain and Ingmar H Riedel-Kruse. Life-science experiments online: Technological frameworks and educational use cases. *Cyber-Physical Laboratories in Engineering and Science Education*, pages 271–304, 2018.

Rebecca C Jordan, Heidi L Ballard, and Tina B Phillips. Key issues and new approaches for evaluating citizen-science learning outcomes. *Frontiers in Ecology and the Environment*, 10(6):307–309, 2012.

B. Kar. Citizen science in risk communication in the era of ict. *Concurrency Computation*, 28(7):2005–2013, 2016.

R. Kermish-Allen, K. Peterman, and C. Bevc. The utility of citizen science projects in k-5 schools: measures of community engagement and student impacts. *Cultural Studies of Science Education*, pages 1–15, 2018.

J.J. Kerski. *Opportunities and challenges in using geospatial technologies for education.* 2015.

Sheri L KLUG, Michelle VIOTTI, Paige GRAFF, Wendy L TAYLOR, Cassie BOWMAN, Laurie ROGERS, and Meg HUFFORD. Participatory exploration: Growing the next generation into science literate citizens through participatory opportunities in mars exploration. In *2010 GSA Denver Annual Meeting*, 2010.

S. Kocaman, B. Anbaroglu, C. Gokceoglu, and O. Altan. A review on citizen science (citsci) applications for disaster management. volume 42, pages 301–306, 2018.

Donna M Kridelbaugh. The use of online citizen-science projects to provide experiential learning opportunities for nonmajor science students. *Journal of microbiology & biology education*, 17(1): 105, 2016.

Erica R Krimmel, Debra L Linton, Travis D Marsico, Anna K Monfils, Ashley B Morris, and Brad R Ruhfel. Collectionseducation. org: Connecting students to citizen science and curated collections. 2017.

Brandon Lawton, Bonnie Eisenhamer, Barbara J. Mattson, and M. Jordan Raddick. Bringing the virtual astronomical observatory to the education community. In JB Jensen, JG Manning, MG Gibbs, and D Daon, editors, *CONNECTING PEOPLE TO SCIENCE: A NATIONAL CONFERENCE ON SCIENCE EDUCATION AND PUBLIC OUTREACH*, volume 457 of *Astronomical Society of the Pacific Conference Series*, pages 283+. Space Telescope Sci Inst; NASA Lunar Sci Inst; NASAs Explorat Program; Infrared Process & Anal Ctr; NASAs Herschel Sci Ctr; Spitzer Sci Ctr; Stratospher Observ Infrared Astron; NASAs Chandra XRay Observ; Univ Chicago Press; Natl Radio Astron Observ; Ball Aerospace; Capitol Coll; Sky Skan; Univ Wyoming CAPER Team; Amer Astron Soc; AAS Educ Off; Solar Dynam Observ; Seiler Instrument; Explore Sci; Pratt St Ale House; MWT Assoc Inc; Amer Elements; Charlesbridge; Celestron, 2012. ISBN 978-1-58381-796-4. 123rd Annual Meeting of the Astronomical-Society-of-the-Pacific, Amer Geophys Union, Baltimore, MD, JUL 30-AUG 03, 2011.

A. Longo and M. Zappatore. A microservice-based mool in acoustics addressing the learning-at-scale scenario. In S Reisman, SI Ahamed, C Demartini, T Conte, L Liu, W Claycomb, M Nakamura, E Tovar, S Cimato, CH Lung, H Takakura, JJ Yang, T Akiyama, Z Zhang, and K Hasan, editors, *2017 IEEE 41ST ANNUAL COMPUTER SOFTWARE AND APPLICATIONS CONFERENCE (COMPSAC), VOL 1*, volume 1 of *Proceedings International Computer Software and Applications Conference*, pages 391–400. IEEE; IEEE Comp Soc, 2017. ISBN 978-1-5386-0367-3. 41st IEEE Annual Computer Software and Applications Conference (COMPSAC), Torino, ITALY, JUL 04-08, 2017.

Amy Lorenz. The influence of a citizen science project: Student attitudes, sense of place, and understanding of science practices. 2016.

E. Lostal Lanza, F. Serrano Sanz, J. A. Carrodeguas Villar, P. Martinez Alonso, F. Sanz Garcia, and C. Val Gascon. Cell images analysis as a case of citizen science for advanced education: Laboratory and school, back and forth. In LG Chova, AL Martinez, and IC Torres, editors, *7TH INTERNATIONAL TECHNOLOGY, EDUCATION AND DEVELOPMENT CONFERENCE (INTED2013)*, INTED Proceedings, pages 2489–2496, 2013. ISBN 978-84-616-2661-8. 7th International Technology, Education and Development Conference (INTED), Valencia, SPAIN, MAR 04-06, 2013.

Maria Luisa de Lazaro, Rafael De Miguel, and Maria Jesus Gonzalez. The real sociedad geografica, a european digital-earth excellence centre: A network for informal learning processes. In LG Chova, AL Martinez, and IC Torres, editors, *ICERI2014: 7TH INTERNATIONAL CONFERENCE OF EDUCATION, RESEARCH AND INNOVATION*, ICERI Proceedings, pages 107–116, 2014. ISBN 978-84-617-2484-0. 7th International Conference of Education, Research and Innovation (ICERI), Seville, SPAIN, NOV 17-19, 2014.

Caitlin K Martin and Brigid Barron. Navigating online learning environments in the classroom. volume 2, pages 307–308. International Society of the Learning Sciences, 2013.

Caitlin K Martin, Denise Nacu, and Nichole Pinkard. An approach to using log data to understand and support 21st century learning activity in k-12 blended learning environments. *Journal of Learning Analytics*, 3(2):37–87, 2016a.

Caitlin K Martin, Denise Nacu, and Nichole Pinkard. Revealing opportunities for 21st century learning: An approach to interpreting user trace log data. *Journal of Learning Analytics*, 3(2): 37–87, 2016b.

C.K. Martin, B. Barron, and V. Mertl. Design methods to study learning across networked systems, co-located spaces, and time. volume 2, pages 309–310, 2013.

K. Masters, E.Y. Oh, J. Cox, B. Simmons, C. Lintott, G. Graham, A. Greenhill, and K. Holmes. Science learning via participation in online citizen science. *Journal of Science Communication*, 15(3), 2016. ISSN 1824-2049.

Sabrina McCormick. After the cap: Risk assessment, citizen science and disaster recovery. *Ecology and Society*, 17(4), 2012. ISSN 17083087.

K. Michalak. Online localization of zooniverse citizen science projects - on the use of translation platforms as tools for translator education. *Teaching English with Technology*, 15(3):61–70, 2015.

N. Mitchell, M. Triska, A. Liberatore, L. Ashcroft, R. Weatherill, and N. Longnecker. Benefits and challenges of incorporating citizen science into university education. *PLoS ONE*, 12(11), nov 1 2017. ISSN 1932-6203.

György Molnár and Zoltán Szűts. Crowdsourcing project as part of nonformal education. pages 943–953, 2018.

Anna K. Monfils, Karen E. Powers, Christopher J. Marshall, Christopher T. Martine, James F. Smith, and L. Alan Prather. Natural history collections: Teaching about biodiversity across time, space, and digital platforms. *SOUTHEASTERN NATURALIST*, 16(10):47–57, September 2017. ISSN 1528-7092.

G. Newman, A. Crall, M. Laituri, J. Graham, T. Stohlgren, J.C. Moore, K. Kodrich, and K.A. Holfelder. Teaching citizen science skills online: Implications for invasive species training programs. *Applied Environmental Education and Communication*, 9(4):276–286, 2010.

Vineet Pandey, Amnon Amir, Justine Debelius, Embriette R Hyde, Tomasz Kosciolek, Rob Knight, and Scott Klemmer. Gut instinct: Creating scientific theories with online learners. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, volume 2017-May, pages 6825–6836. ACM, 2017. ISBN 978-1-4503-4655-9. ACM SIGCHI Conference on Human Factors in Computing Systems (CHI), Denver, CO, MAY 06-11, 2017.

Charles Aaron Price. *Scientific literacy of adult participants in an online citizen science project*. Tufts University, 2011.

Jordan Raddick, K Carney, LF Fortson, PL Gay, CJ Lintott, and Galaxy Zoo Team. How can citizen science participation improve science learning? In *Bulletin of the American Astronomical Society*, volume 41, page 422, 2009.

Harsh R Shah and Luis R Martinez. Current approaches in implementing citizen science in the classroom. *Journal of microbiology & biology education*, 17(1): 17, 2016.

A. Striner and J. Preece. Streambed: Training citizen scientists to make qualitative judgments using embodied virtual reality training. volume 07-12-May-2016, pages 1201–1207, 2016.

Nancy Trautmann, Jennifer Fee, and Phil Kahler. Flying into inquiry. *The Science Teacher*, 79(9):45, 2012.

Nancy M Trautmann. *Citizen science: 15 lessons that bring biology to life, 6-12*. NSTA Press, 2013.

Ginger Tsueng, Arun Kumar, Max Nanis, and Andrew I Su. Aligning interests: Integrating citizen science efforts into schools through service requirements. *bioRxiv*, page 304766, 2018.

David A Ucko. Science centers in a new world of learning. *Curator: The Museum Journal*, 56(1):21–30, 2013.

Margaret A. Voss and Caren B. Cooper. Using a free online citizen-science project to teach observation & quantification of animal behavior. *The American Biology Teacher*, 72(7):437–443, September 2010. ISSN 00027685, 19384211.

Duane Wallace and Alec Bodzin. Developing scientific citizenship identity using mobile learning and authentic practice. *Electronic Journal of Science Education*, 21(6), 2017.

Ginny Webb. A review of microbiology service learning. *FEMS microbiology letters*, 364(4):fnx032, 2017.

A.2 Irrelevant Citizen Science Topics

These papers concerned relevant projects, but topics of irrelevance to this thesis:

Stuart Allan and Joanna Redden. Making citizen science newsworthy in the era of big data. *JCOM: Journal of Science Communication*, 16(2):1i–1i, 2017.

K. De Cocker, S.F.M. Chastin, I. De Bourdeaudhuij, I. Imbo, J. Stragier, and G. Cardon. Citizen science to communicate about public health messages: The reach of a playful online survey on sitting time and physical activity. *Health Communication*, pages 1–6, 2018.

A. Fresa, B. Justrel, V. Bachi, and N. Forbes. Civic epistemologies - development of a roadmap for citizen researchers in the age of digital culture. pages 8–14, 2015.

Yaela N Golumbic, Daniela Orr, Ayelet Baram-Tsabari, and Barak Fishbain. Between vision and reality: A study of scientists' views on citizen science. *Citizen Science: Theory and Practice*, 2 (1), 2017.

Johann Höchtl, Judith Schossböck, Thomas J Lampoltshammer, and Peter Parycek. The citizen scientist in the epolicy cycle. In *Government 3.0–Next Generation Government Technology Infrastructure and Services*, pages 37–61. Springer, 2017.

MW Home. Diy zooniverse citizen science project: engaging the public with your museum's collections and data.

Philip J Marshall, Chris J Lintott, and Leigh N Fletcher. Ideas for citizen science in astronomy. *Annual Review of Astronomy and Astrophysics*, 53:247–278, 2015. ISSN 0066-4146.

BJH Méndez, B Day, PL Gay, SH Jacoby, MJ Raddick, CE Walker, and SM Pompea. The spectrum of citizen science projects in astronomy and space science. In *Science Education and Outreach: Forging a Path to the Future*, volume 431, page 324, 2010.

Oleguer Sagarra, Mario Gutiérrez-Roig, Isabelle Bonhoure, and Josep Perelló. Citizen science practices for computational social science research: The conceptualization of pop-up experiments. *Frontiers in physics*, 3:93, 2016.

J Patrick Woolley, Michelle L McGowan, Harriet JA Teare, Victoria Coathup, Jennifer R Fishman, Richard A Settersten, Sigrid Sterckx, Jane Kaye, and Eric T Juengst. Citizen science or scientific citizenship? disentangling the uses of public engagement rhetoric in national research initiatives. *BMC medical ethics*, 17(1):33, 2016.

A.3 Irrelevant - Not Citizen Science

These papers were of no relevance to Citizen Science:

V. Ãzdemir, K.F. Badr, E.S. Dove, L. Endrenyi, C.J. Geraci, P.J. Hotez, D. Milius, M. Neves-Pereira, T. Pang, C.N. Rotimi, R. Sabra, C.N. Sarkissian, S. Srivastava, H. Tims, N.K. Zgheib, and I. Kickbusch. Crowd-funded micro-grants for genomics and big data: An actionable idea connecting small (artisan) science, infrastructure science, and citizen philanthropy. *OMICS A Journal of Integrative Biology*, 17(4): 161–172, April 2013. ISSN 1536-2310.

Kuku Joseph Aduku, Mike Thelwall, and Kayvan Kousha. Do mendeley reader counts reflect the scholarly impact of conference papers? an investigation of computer science and engineering. *Scientometrics*, 112(1):

573–581, 2017.

Tanja Aitamurto. Benefits, challenges, and value creation. *The Routledge Companion to Digital Journalism Studies*, page 185, 2016.

Emmanuel Akyeampong. Indigenous knowledge and maritime fishing in west africa: the case of ghana. *Tribes and Tribals*, pages 173–182, 2007.

Daniel P. Aldrich. Rethinking civil society-state relations in japan after the fukushima accident. *POLITY*, 45(2): 249–264, April 2013. ISSN 0032-3497.

Sue Allen and Joshua P Gutwill. Creating a program to deepen family inquiry at interactive science exhibits. *Curator: The Museum Journal*, 52(3):289–306, 2009.

Danielle Alvarez, Connie Kang, Denise Lin, June Tran, and Tiffany Wu. Developing biodiversity indicators for los angeles county.

LENY CADIZ AMPARO. Factors influencing uplb scientists in studying climate change. 2013.

Katherine L. Anderson, Ute Kaden, Patrick S. Druckenmiller, Sarah Fowell, Mark A. Spangler, Falk Huettmann, and Stefanie M. Ickert-Bond. Arctic science education using public museum collections from the university of alaska museum: an evolving and expanding landscape. *ARCTIC SCIENCE*, 3(3, SI):635–653, September 2017. ISSN 2368-7460.

T. Anderson. Open access scholarly publications as oer. *International Review of Research in Open and Distance Learning*, 14(2):81–95, 2013. ISSN 1492-3831.

Clinton J Andrews. Communicative science. In *Joint Fact-Finding in Urban Planning and Environmental Disputes*, pages 76–85. Routledge, 2016.

Fátima Antunes and Paula Guimarães. Recognition of prior learning: valuing, learning through transitions for individual and collective purposes. In *Changing configurations of adult education in transitional times: conference proceedings [of the 7th European Research Conference]*, pages 201–212. Esrea, 2013.

SPEEDS ARE. Connextion. 2017.

Emílio José Montero Arruda Filho and Igor de Jesus Lobato Pompeu Gammarano. For every 'game over' there is a 'play again': Analysis of user preferences regarding 7th-and 8th-generation video games consoles. *The Journal of High Technology Management Research*, 2018.

Larissa L Bailey, Darryl I MacKenzie, and James D Nichols. Advances and applications of occupancy models. *Methods in Ecology and Evolution*, 5(12): 1269–1279, 2014.

Fraser Baker, Claire L Smith, and Gina Cavan. Remote sensing. 1965.

Mara Balestrini, Tomas Diez, Paul Marshall, Alex Gluhak, and Yvonne Rogers. Iot community technologies: leaving users to their own devices or orchestration of engagement? *EAI Endorsed Transactions on Internet of Things*, 1(1), 2015.

Brigid Barron. Learning across setting and time: Catalysts for synergy. In *BJEP Monograph Series II, Number 11-Learning Beyond the Classroom*, volume 7, pages 7–21. British Psychological Society, 2015.

Andreas Baur. Crowdsourced formal verification: a business case analysis toward a human-centered business model. Technical report, NAVAL POSTGRADUATE SCHOOL MONTEREY CA, 2015.

Petra Benyei, Nerea Turreira-Garcia, Martí Orta-Martínez, and Mar Cartró-Sabaté. Globalized conflicts, globalized responses. changing manners of contestation among indigenous communities. In *Hunter-gatherers in a Changing World*, pages 233–250. Springer, 2017.

E. Bertin, R. Pillay, and C. Marmo. Web-based visualization of very large scientific astronomy imagery. *Astronomy and Computing*, 10:43–53, 2015.

A. Bevan, X. Li, M. MartinÃn Torres, S. Green, Y. Xia, K. Zhao, Z. Zhao, S. Ma, W. Cao, and T. Rehren. Computer vision, archaeological classification and china's terracotta warriors. *Journal of Archaeological Science*, 49(1):249–254, September 2014. ISSN 0305-4403.

B. Boccardi, M. Fragona, and G. Parolini. Popularisation of physics in the wild. volume 145, pages 609–615, 2014.

V. Bocci, G. Chiodi, P. Fresch, F. Iacoangeli, and L. Recchia. An educational distributed cosmic ray detector network based on ardusipm. volume 898, 2017.

Martin Bogner. Conception and implementation of a collaborative data science platform.

II Bohannon and R Richard. Habitat fragmentation in western north dakota after the introduction of hydraulic fracturing. 2017.

Frank A Bosco, Krista L Uggerslev, and Piers Steel. metabus as a vehicle for facilitating meta-analysis. *Human Resource Management Review*, 27(1): 237–254, 2017.

Sue Bowler. Looking forward for geophysics. *Astronomy & Geophysics*, 58(3): 3–27, 2017.

Doreen S Boyd, Bethany Jackson, Jessica Wardlaw, Giles M Foody, Stuart Marsh, and Kevin Bales. Slavery from space: Demonstrating the role for satellite remote sensing to inform evidence-based action related to un sdg number 8. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2018.

S. Broadbent and F. Cara. Seeking control in a precarious environment: Sustainable practices as an adaptive strategy to living under uncertainty. *Sustainability (Switzerland)*, 10(5), May 2018. ISSN 2071-1050.

Katie L Burke. Interface facts. *American Scientist*, 100(6):463, 2012.

G. Burnett and P. Moderator. Information science approaches to studying virtual organizations: A panel. *Proceedings of the ASIST Annual Meeting*, 47, 2010.

John Stephen Chalk. Gamefully designed language learning an alternate reality for efl learning environments.

Marjorie Chan and Julia Kahmann-Robinson. Mars for earthlings: An analog approach to mars in undergraduate education. *Astrobiology*, 14(1):42–49, 2014.

A. Collier-Oxandale, E. Coffey, J. Thorson, J. Johnston, and M. Hannigan. Comparing building and neighborhood-scale variability of co2 and o3 to inform deployment considerations for low-cost sensor system use. *Sensors (Switzerland)*, 18(5), 2018.

Carly N Cook, Bonnie C Wintle, Stephen C Aldrich, and Brendan A Wintle. Using strategic foresight to assess conservation opportunity. *Conservation biology*, 28(6):1474–1483, 2014.

Joana Costa, Catarina Silva, Mário Antunes, and Bernardete Ribeiro. Get your jokes right: ask the crowd. In *International Conference on Model and Data Engineering*, pages 178–185. Springer, 2011.

Vickie Curtis. Evaluating the motivations and expectations of those attending a public astronomy event. *Communicating Astronomy to the Public*, 13:14–18, 2013.

Vickie Curtis. Public engagement through the development of science-based computer games: The wellcome trust "gamify your phd" initiative. *Science Communication*, 36(3):379–387, 2014.

Anna R. Davies, Ferne Edwards, Brigida Marovelli, Oona Morrow, Monika Rut, and Marion Weymes. Creative construction: crafting, negotiating and performing urban food sharing landscapes. *AREA*, 49(4):510–518, December 2017. ISSN 0004-0894.

D.G. De Paor. Virtual rocks. *GSA Today*, 26(8):4–11, 2016.

František Dinnbier. Propagating star formation. 2017.

M. DolejÅovÃ, D. Kera, C. Storni, R.A. Khot, I.J. Clement, I. Pavelka, and P. Kishor. Digital health & self-experimentation: Design challenges & provocations. volume Part F127655, pages 510–517, 2017.

L.J. Dorward, J.C. Mittermeier, C. Sandbrook, and F. Spooner. PokÃmon go: Benefits, costs, and lessons for the conservation movement. *Conservation Letters*, 10(1):160–165, January 2017. ISSN 1755-263X.

Steve Dudley and Jennifer Smart. How social are ornithologists? *Ibis*, 158(4): 894–898, 2016.

S.W. Duxbury. Information creation on online drug forums: How drug use becomes moral on the margins of science. *Current Sociology*, 66(3):431–448, May 2018. ISSN 0011-3921.

S. Ellis. A history of collaboration, a future in crowdsourcing: Positive impacts of cooperation on british librarianship. *Libri*, 64(1):1–10, March 2014. ISSN 0024-2667.

Dmitry Epstein and Gilly Leshed. The magic sauce: Practices of facilitation in online policy deliberation. *Journal of Public Deliberation*, 12(1):4, 2016.

S. Faridani, B. Lee, S. Glasscock, J. Rappole, D. Song, and K. Goldberg. A networked telerobotic observatory for collaborative remote observation of avian activity and range change. pages 56–61, 2009.

Piyum Fernando, Matthew Pandelakis, and Stacey Kuznetsov. Practicing diybiology in an hci setting. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, pages 2064–2071. ACM, 2016.

Meghan Ferriter. Inviting engagement, supporting success. *Collections Vol 12*, (2):97, 2016.

Casey Fiesler, Shannon Morrison, R Benjamin Shapiro, and Amy S Bruckman. Growing their own: Legitimate peripheral participation for computational learning in an online fandom community. In *CSCW*, pages 1375–1386, 2017.

Gordon Fletcher, Anita Greenhill, Marie

Griffiths, Kate Holmes, and Rachel McLean. Creatively prototyping the future high street. *Production Planning & Control*, 27(6):477–489, 2016.

Ashleigh Frederiksen, Robert Andrew McLeman, and Tim L Elcombe. Building backyard ice rinks in canada: an exploratory study. *Leisure/Loisir*, 42(1): 47–68, 2018.

Bruno S Frey and Reto Jegen. Motivation crowding theory. *Journal of economic surveys*, 15(5):589–611, 2001.

Carlos Galan-Diaz, Peter Edwards, John D. Nelson, and Rene van der Wal. Digital innovation through partnership between nature conservation organisations and academia: A qualitative impact assessment. *AMBIO*, 44(4, SI): S538–S549, November 2015. ISSN 0044-7447.

A. Gesser-Edelsburg, Y. Shir-Raz, S. Hayek, and O. Sassoni-Bar Lev. What does the public know about ebola? the public's risk perceptions regarding the current ebola outbreak in an as-yet unaffected country. *American Journal of Infection Control*, 43(7):669–675, July 2015. ISSN 0196-6553.

Silvia Gherardi and Francesco Miele. Knowledge management from a social perspective: The contribution of practice-based studies. In *The Palgrave Handbook of Knowledge Management*, pages 151–176. Springer, 2018.

Kathryn Gillespie and Rosemary-Claire Collard. *Critical animal geographies: Politics, intersections and hierarchies in a multispecies world.* Routledge, 2015.

H. GoÃau, P. Bonnet, J. Barbe, V. Bakic, A. Joly, and J.-F. Molino. Multi-organ plant identification. pages 41–44, 2012.

Edward L Gomez and Michael T Fitzgerald. Robotic telescopes in education. *Astronomical Review*, 13(1): 28–68, 2017.

E.L. Gomez. Las cumbres observatory: Building a global telescope network from the ground up. *Proceedings of the International Astronomical Union*, 10 (H16):646, 2012.

Edit Görögh, Michela Vignoli, Stephan Gauch, Clemens Blümel, Peter Kraker, Ilire Hasani-Mavriqi, Daniela Luzi, Mappet Walker, Eleni Toli, and Electra Sifacaki. Opening up new channels for scholarly review, dissemination, and assessment. In *Proceedings of the 13th International Symposium on Open Collaboration*, page 6. ACM, 2017.

S.J.J. Gould, S. Wiseman, D. Furniss, I. Iacovides, C.I. Jennett, and A.L. Cox. Moods: Building massive open online diaries for researchers, teachers and contributors. pages 2281–2286, 2014.

C. Gura, E. Wandl-Vogt, A. Dorn, A. Losada, and A. Benito. Co-designing innovation networks for cross-sectoral collaboration on the example of exploreat! volume Part F132203, 2017.

Louis Alberto Gutierrez. *Noise reduction in user generated datasets.* PhD thesis, Rensselaer Polytechnic Institute, 2014.

Myron P Gutmann, Emily Klancher Merchant, and Evan Roberts. 'big data' in economic history. *The journal of economic history*, 78(1):268–299, 2018.

R.J. Hansen and E.M. Brady. Research in the osher lifelong learning institute network. *Journal of Continuing Higher Education*, 61(3):143–150, 2013.

Mahboobeh Harandi, Kevin Crowston, and Corey Jackson. Perceptions of machine learning.

Benjamin K Haywood. A 'sense of place' in public participation in scientific research. *Science education*, 98(1):64–83, 2014.

O. Hazzan and C.A. Shaffer. Big data in computer science education research. pages 591–592, 2015.

Daqing He and Wei Jeng. Scholarly collaboration on the academic social web. *Synthesis Lectures on Information Concepts, Retrieval, and Services*, 8(1): 1–106, 2016.

Y. He, J. Preece, J. Hammock, B. Butler, and D. Pauw. Understanding data providers in a global scientific data hub. volume 2015-January, pages 215–218, 2015.

Robert Heck, Jacob F Sherson, et al. Are strategies in physics discrete? a remote controlled investigation. In *APS Division of Atomic, Molecular and Optical Physics Meeting Abstracts*, 2017.

J.E. Herbert-Read, M. Romenskyy, and D.J.T. Sumpter. A turing test for collective motion. *Biology Letters*, 11(12), December 2015. ISSN 1744-9561.

Joan Herman. Opportunity to learn. *Encyclopedia of Science Education*, pages 725–727, 2015.

Ingeborg Heyer, Stephanie J Slater, and Timothy F Slater. *Establishing the empirical relationship between non-science majoring undergraduate learners' spatial thinking skills and their conceptual astronomy knowledge.* University of Wyoming Laramie, WY, 2012.

E.-L.S. Hinckley, S.P. Anderson, J.S. Baron, P.D. Blanken, G.B. Bonan, W.D. Bowman, S.C. Elmendorf, N. Fierer, A.M. Fox, K.J. Goodman, K.D. Jones, D.L. Lombardozzi, C.K. Lunch, J.C. Neff, M.D. SanClements, K.N. Suding, and W.R. Wieder. Optimizing available network resources to address questions in environmental biogeochemistry. *BioScience*, 66(4):317–326, April 2016. ISSN 0006-3568.

Nina ST Hirata. Galaxy image classification.

Angelique Greer Hjarding. Are biodiversity data gaps in urban areas a social justice issue? In *TDWG 2014 ANNUAL CONFERENCE*, 2014.

Elizabeth A Hobson, Grace Smith-Vidaurre, and Alejandro Salinas-Melgoza. History of nonnative monk parakeets in mexico. *PloS one*, 12(9): e0184771, 2017.

C. Hodapp, M. Robbins, J. Gray, and A. Graettinger. Collecting damage information after natural disasters. 2013.

Richard Holliman and Vickie Curtis. Online media. *Encyclopedia of science education*, pages 718–725, 2015.

Iris Holmes. Checking in with jamaica's endangered frogs. *FrogLog*, page 38.

Kate Holmes, Anita Greenhill, and Rachel McLean. Creating communities: the use of technology in craft and diy communities of practice. *Journal of Systems and Information Technology*, 16 (4):277–295, 2014.

Julie S Hui and Elizabeth M Gerber. Crowdfunding science: Sharing research with an extended audience. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*, pages 31–43. ACM, 2015.

Sam Illingworth and Andreas Prokop. Science communication in the field of fundamental biomedical research. In *Seminars in cell & developmental biology.* Elsevier, 2017.

Rob Inkpen. New technologies and the political economy of geomorphology. *The Canadian Geographer/Le Géographe canadien*, 2018.

Nick JB Isaac and Michael JO Pocock. Bias and information in biological records. *Biological Journal of the Linnean Society*, 115(3):522–531, 2015.

Barash Carol Isaacson. Global genomic knowledge sharing – a call for affirmative action, 2014.

M. Isenburg. Open data and its benefits to society: Remembering the buchenwald concentration camp. volume 2017-October, 2017.

H.R. Jamali, D. Nicholas, and E. Herman. Scholarly reputation in the digital age and the role of emerging platforms and mechanisms. *Research Evaluation*, 25(1):37–49, January 2016. ISSN 0958-2029.

Paul-Ariel Kenigsberg, Jean-Pierre Aquino, Alain Bérard, François Brémond, Kevin Charras, Tom Dening, Rose-Marie Droës, Fabrice Gzil, Ben Hicks, Anthea Innes, et al. Assistive technologies to address capabilities of people with dementia: from research to practice. *Dementia*, page 1471301217714093, 2017.

Z. Khan, S.L. Kiani, and K. Soomro. A framework for cloud-based context-aware information services for citizens in smart cities. *Metallography, Microstructure, and Analysis*, 3(1), 2014.

Patricia Kristjanson, Elizabeth Bryan, Quinn Bernier, Jennifer Twyman, Ruth Meinzen-Dick, Caitlin Kieran, Claudia Ringler, Christine Jost, and Cheryl Doss. Addressing gender in agricultural research for development in the face of a changing climate: where are we and where should we be going? *International Journal of Agricultural Sustainability*, 15 (5):482–500, 2017.

ALANA KRUG-MACLEOD. Motivating and mobilizing youth: Conservation awareness-building in antarctica, the backyard, and online. In *AND ENDANGERED SPECIES CONFERENCE*, page 212, 2016.

S. Kuznetsov, C.J. Santana, and E. Long. Everyday food science as a design space for community literacy and habitual sustainable practice. In *34TH ANNUAL CHI CONFERENCE ON HUMAN FACTORS IN COMPUTING SYSTEMS, CHI 2016*, pages 1786–1797. Assoc Comp Machinery; SIG CHI, 2016. ISBN 978-1-4503-3362-7. 34th Annual CHI Conference on Human Factors in Computing Systems (CHI4GOOD), San Jose, CA, MAY 07-12, 2016.

Roberta Kwok. Historical data: Hidden in the past. *Nature*, 549(7672):419–421, 2017.

Eleni A Kyza. Online inquiry environments. *Encyclopedia of Science Education*, pages 715–718, 2015.

Lesley A Langa. Does twitter help museums engage with visitors? *IConference 2014 Proceedings*, 2014.

Colin Latchem. Learning technology and lifelong informal, self-directed, and non-formal learning. *The Wiley Handbook of Learning Technology*, 1:180, 2016.

Kristin Fontichiaro Amy Lennex. Data literacy in the real world.

Nicola Lettieri, Antonio Altamura,

Rosalba Giugno, Alfonso Guarino, Delfina Malandrino, Alfredo Pulvirenti, Francesco Vicidomini, and Rocco Zaccagnino. Ex machina: Analytical platforms, law and the challenges of computational legal science. *Future Internet*, 10(5):37, 2018.

Roman Lukyanenko and Jeffrey Parsons. Conceptual modeling of open information systems.

Roman Lukyanenko and Jeffrey Parsons. Rethinking data quality as an outcome of conceptual modeling choices. In *16th International Conference on Information Quality*, pages 1–16, 2011.

Roman Lukyanenko, Jeffrey Parsons, and Binny Samuel. Artifact sampling: Using multiple information technology artifacts to increase research rigor. In *Proceedings of the 51st Hawaii International Conference on System Sciences*, 2018.

E. B. Mackay, M. E. Wilkinson, C. J. A. Macleod, K. Beven, B. J. Percy, M. G. Macklin, P. F. Quinn, M. Stutter, and P. M. Haygarth. Digital catchment observatories: A platform for engagement and knowledge exchange between catchment scientists, policy makers, and local communities. *WATER RESOURCES RESEARCH*, 51(6):4815–4822, June 2015. ISSN 0043-1397.

Philip Macnaghten. The metis of responsible innovation.

Alexia Maddox and Linlin Zhao. University library strategy development: A conceptual model of researcher performance to inform service delivery. *New Review of Academic Librarianship*, 23 (2-3):125–135, 2017.

LAURA W MARTIN. Free-choice learning: What does it mean? In *Intersections of Formal and Informal Science*, pages 46–58. Routledge, 2016.

Andrew May and Tracy Ross. The design of civic technology: factors that influence public participation and impact. *Ergonomics*, 61(2):214–225, 2018.

Pauric McGowan, Richard Blundel, and Kristen Reid. Delivering effective enterprise education–the role of learning design and technology. 2014.

G.R. McKercher, J.A. Salmond, and J.K. Vanos. Characteristics and applications of small, portable gaseous air pollution monitors. *Environmental Pollution*, 223: 102–110, April 2017. ISSN 0269-7491.

EA McKinnon, C Artuso, and OP Love. The mystery of the missing warbler. *Ecology*, 98(7):1970–1972, 2017.

John McNutt, Chao Guo, Lauri Goldkind, and Seongho An. Technology in nonprofit organizations and voluntary action. 2018.

R.V.D. Meer and G. CimÃ. Opportunities within asterics. In A Capone, G DeBonis, I DiPalma, and C Perrina, editors, *VERY LARGE VOLUME NEUTRINO TELESCOPE (VLVNT-2015)*, volume 116 of *EPJ Web of Conferences*. INFN, 2016. ISBN 978-2-7598-2016-0. 7th Biannual Very Large Volume Neutrino Telescope Workshop (VLVnT), La Sapienza Univ, Dept Phys, Roma, ITALY, SEP 14-16, 2015.

S.R. Mekaru and J.S. Brownstein. One health in social networks and social media. *OIE Revue Scientifique et Technique*, 33(2):629–637, August 2014. ISSN 0253-1933.

Germán M Mendoza-Silva, Luis E Rodríguez-Pupo, Joaquín Torres-Sospedra, and Joaquín Huerta-Guijarro. Solutions for signal mapping campaigns of wi-fi networks.

Ines Mergel. Big data in public affairs education. *Journal of Public Affairs Education*, 22(2):231–248, 2016.

F. Metze, E. Riebling, A.S. Warlaumont, and E. Bergelson. Virtual machines and containers as a platform for experimentation. In *17TH ANNUAL CONFERENCE OF THE INTERNATIONAL SPEECH COMMUNICATION ASSOCIATION (INTERSPEECH 2016), VOLS 1-5: UNDERSTANDING SPEECH PROCESSING IN HUMANS AND MACHINES*, volume 08-12-September-2016 of *Interspeech*, pages 1603–1607. apple; amazon alexa; Google; Microsoft; ebay; facebook; YAHOO JAPAN; Baidu Res; IBM Res; CIRRUS LOGIC; DATATANG; NUANCE; Speechocean Ltd; Yandex; Raytheon Technol, 2016. ISBN 978-1-5108-3313-5. 17th Annual Conference of the International-Speech-Communication-Association (INTERSPEECH 2016), San Francisco, CA, SEP 08-12, 2016.

C. Michel, A.T. Woods, M. NeuhÃuser, A. Landgraf, and C. Spence. Rotating plates: Online study demonstrates the importance of orientation in the plating of food. *Food Quality and Preference*, 44:194–202, September 2015. ISSN 0950-3293.

Pietro Michelucci. Human computation and convergence. *Handbook of science and technology convergence*, pages 455–474, 2016.

Linda Morrice. Issues of recognition and participation in changing times: the inclusion of refugees in higher education in the uk. In *7th European Research Conference, 4-7 September 2013, Humboldt-University, Berlin*, pages 135–143. European Society for Research on the Education of Adult, 2013.

Karyn Morrissey, Peter Kinderman, Eleanor Pontin, Sara Tai, and Mathias Schwannauer. Web based health surveys: Using a two step heckman model to examine their potential for population health analysis. *Social Science & Medicine*, 163:45–53, 2016.

Caley Mulholland. Reading the landscape for ecological restoration: a resource to develop students' site assessment practice. 2016.

Shinnosuke Nakayama, Manuel Ruiz Marin, Maximo Camacho, and Maurizio Porfiri. Plasticity in leader-follower roles in human teams. *SCIENTIFIC REPORTS*, 7, November 2017. ISSN 2045-2322.

Beth Simone Noveck. *Smart citizens, smarter state: The technologies of expertise and the future of governing*. Harvard University Press, 2015.

P. øAkoda. Astroinformatics: Getting new knowledge from the astronomical data avalanche. *Advances in Intelligent Systems and Computing*, 210:15, 2013.

Erinma Ochu. 11 in search of lost purpose. *Doing Research In and On the Digital: Research Methods across Fields of Inquiry*, 2018.

Pier Oddone. Cerncourier.

G. Paine. Ecologies of listening and presence: Perspectives from a practitioner. *Contemporary Music Review*, 35 (3):362–371, June 2016. ISSN 0749-4467.

Maria Victoria Palacin-Silva. Understanding civic participation in environmental sensing: A values driven approach. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*, page DC16. ACM, 2018.

Andrew J Pearl and Robert K Christensen. First-year student motivations for service-learning: An application of the volunteer functions inventory. *Michigan Journal of Community Service Learning*, 23(2):66–82, 2017.

Elissa Pearson, Hayley Tindle, Monika Ferguson, Jillian Ryan, and Carla Litchfield. Can we tweet, post, and share our way to a more sustainable society? a review of the current contributions and future potential of# socialmediaforsustainability. *Annual Review of Environment and Resources*, 41:363–397, 2016.

M.K. Pedersen, B. Skyum, R. Heck, R. MÃller, M. Bason, A. Lieberoth, and J.F. Sherson. Virtual learning environment for interactive engagement with advanced quantum mechanics. *Physical Review Physics Education Research*, 12 (1), April 2016. ISSN 2469-9896.

John H Phan, Sonal Kothari, and May D Wang. omniclassifier: a desktop grid computing system for big data prediction modeling. In *Proceedings of the 5th ACM Conference on Bioinformatics, Computational Biology, and Health Informatics*, pages 514–523. ACM, 2014.

Lotte Philipsen and Rikke Schmidt Kjærgaard. Scientific storytelling: Visualizing for public audiences dj ukev el dh u is. In *The Aesthetics of Scientific Data Representation*, pages 109–122. Routledge, 2017.

N. Pitrelli. Big data and digital methods in science communication research: Opportunities, challenges and limits. *Journal of Science Communication*, 16(2), 2017.

N.F. Polys, P. Sforza, W. Cully Hession, and J. Munsell. Extensible experiences: Fusality for stream and field. pages 179–180, 2016.

Philip Pond and Jeff Lewis. Riots and twitter: connective politics, social media and framing discourses in the digital public sphere. *Information, Communication & Society*, pages 1–19, 2017.

C.P. Pritscher. *Learning what to ignore: Connecting multidiscipline content and process*. 2013.

Andreas Prokop and Sam Illingworth. Aiming for long-term, objective-driven science communication in the uk. *F1000Research*, 5, 2016.

FRANCISCO QUEIROZ and REJANE SPITZ. Design management and professional practice. 2016.

Francisco Queiroz and Rejane Spitz. Position paper: Collaborative gamification design for scientific software. 2016.

Mia Ridge. Citizen history and its discontents. 2014.

J.C. Rimland, D. Coughlin, D.L. Hall, and J.L. Graham. Advances in data representation for hard/soft information fusion. volume 8407, 2012.

Mihail C Roco. Principles and methods that facilitate convergence. *Handbook of Science and Technology Convergence*, pages 1–20, 2014.

Mihail C Roco, George Whitesides, Jim Murday, Placid M Ferreira, Giorgio Ascoli, Chin Hua Kong, Clayton Teague, Roop Mahajan, David Rejeski, Eli Yablonovitch, et al. Methods to improve and expedite convergence. In *Convergence of Knowledge, Technology and Society*, pages 139–184. Springer, 2013.

Maria A. Rodriguez and Jose M. Espinoza. Development of a teaching methodology for undergraduate human development in psychology. *PROPOSITOS Y REPRESENTACIONES*, 3(1): 71–97, January 2015. ISSN 2307-7999.

Hillary Rosner. Data on wings. *Scientific American*, 308(2):68–73, 2013.

A. A. Royem, C. K. Mui, D. R. Fuka, and M. T. Walter. Proposing a low-tech, affordable, accurate stream stage monitoring system. *TRANSACTIONS OF THE ASABE*, 55(6):2237–2242, November 2012. ISSN 2151-0032.

Johanna M Russ. Comparing evaluation methodologies for a digital exhibition: The end of tobacco road: Scenes from liggett & myers tobacco company's final days in durham, north carolina, 1999. 2008.

Alan Salvado Romero, Manel Jimenez-Morales, and Carolina Sourdis. The interactive documentary as a creative and artistic experience for youth empowerment: The hebe webdoc as a case of study. *PEDAGOGIA SOCIAL REVISTA INTERUNIVERSITARIA*, (30): 91–104, July 2017. ISSN 1989-9742.

H. Sangster, C. Jones, and N. Macdonald. The co-evolution of historical source materials in the geophysical, hydrological and meteorological sciences: Learning from the past and moving forward. *Progress in Physical Geography*, 42(1):61–82, February 2018. ISSN 0309-1333.

TERESA SCASSA. A guide for researchers and citizen scientists.

H. Scholten, S. Fruijtier, E. Dias, S. Hettinga, M. Opmeer, W.S. van Leeuwen, M. Linde, S. Bos, R. Vaughan, H. van Kaam, N. van Manen, and C. Fruijtier. Geocraft as a means to support the development of smart cities, getting the people of the place involved - youth included -. *Quality Innovation Prosperity*, 21(1):119–150, 2017. ISSN 1335-1745.

B.A. Shaby. The open-faced sandwich adjustment for mcmc using estimating functions. *Journal of Computational and Graphical Statistics*, 23(3):853–876, September 2014. ISSN 1061-8600.

Kalpana Shankar and Kristin Eschenfelder. Organizational and institutional work in data infrastructures. *Proceedings of the Association for Information Science and Technology*, 54(1):595–598, 2017.

Waralak Vongdoiwang Siricharoen and Nattanun Siricharoen. Infographic utility in accelerating better health communication. *Mobile Networks and Applications*, 23(1):57–67, 2018.

David Roy Smith. A walk in the park: Is pokémon go foreshadowing the future of biodiversity research and scientific outreach? *EMBO reports*, page e201643213, 2016.

Joe Smith, George Revill, and Kim Hammond. Voicing climate change? television, public engagement and the politics of voice. *Transactions of the Institute of British Geographers*, 2018.

T. F. Smith and N. S. Lazarow. Social learning and the adaptive management framework. *Journal of Coastal Research*, (39):952–954, win 2006. ISSN 07490208, 15515036.

E.J. Sobo, A. Huhn, A. Sannwald, and L. Thurman. Information curation among vaccine cautious parents: Web 2.0, pinterest thinking, and pediatric vaccination choice. *Medical Anthropology: Cross Cultural Studies in Health and Illness*, 35(6):529–546, 2016. ISSN 0145-9740.

Daniela Soleri, Jonathan W Long, Mónica D Ramirez-Andreotta, Ruth Eitemiller, and Rajul PandyaÇ. Finding pathways to more equitable and meaningful public-scientist partnerships. *Citizen Science: Theory and Practice. 1 (1): 9*, 1(1), 2016.

Hwanseok Song, Jonathon P Schuldt, Poppy L McLeod, Rhiannon L Crain, and Janis L Dickinson. Group norm violations in an online environmental social network: Effects on impression formation and intergroup judgments. *Group Processes & Intergroup Relations*, 21(3): 422–437, April 2018. ISSN 1368-4302.

Patricia A. Soranno, Kendra S. Cheruvelil, Kevin C. Elliott, and Georgina M. Montgomery. It's good to share: Why environmental scientists' ethics are out of date. *BIOSCIENCE*, 65(1):69–73, January 2015. ISSN 0006-3568.

K. Starbird, D. Dailey, A.H. Walker, T.M. Leschine, R. Pavia, and A. Bostrom. Social media, public participation, and the 2010 bp deepwater horizon oil spill. *Human and Ecological Risk Assessment*, 21(3):605–630, April 2015. ISSN 1080-7039.

David Stevens and Kieron O'hara. *The Devil's Long Tail: Religious and Other Radicals in the Internet Marketplace.* Oxford University Press, 2015.

Paul Szerlip. Worldwide infrastructure for neuroevolution: A modular library to turn any evolutionary domain into an online interactive platform. 2015.

Paul Szerlip and Kenneth O Stanley. Steps toward a modular library for turning any evolutionary domain into an online interactive platform. In *H. Sayama, J. Rieffel, S. Risi, R. Doursat, & H. Lipson (Eds.), Artificial life 14: Proceedings of the Fourteenth International Conference on the Synthesis and Simulation of Living Systems*, pages 900–907. Citeseer, 2014.

M. Terras. *Crowdsourcing in the Digital Humanities*. 2015.

Raffaele Trequattrini, Rosa Lombardi, Alessandra Lardo, Sara Della Rosa, and Francesco Bolici. A review of virtual health networks: online communities that manage health information reducing environmental uncertainty. In JC Spender, G Schiuma, and V Albino, editors, *IFKAD 2015: 10TH INTERNATIONAL FORUM ON KNOWLEDGE ASSET DYNAMICS: CULTURE, INNOVATION AND ENTREPRENEURSHIP: CONNECTING THE KNOWLEDGE DOTS*, pages 1504–1514. Inst Knowledge Asset Management; Univ Basilicata; Arts Business Ltd, 2015. ISBN 978-88-96687-07-9. 10th International Forum on Knowledge Asset Dynamics (IFKAD), Polytechn Univ Bari, Bari, ITALY, JUN 10-12, 2015.

Pauric McGowan UU, Richard Blundel OU, and Kristen Reid OU. Delivering effective enterprise education–the role of learning design and technology. *Organization*, 76:90–112.

Emmanuelle Vaast and Hani Safadi. Talking about us and them: The formation of a new occupation's ecosystem with digital technology use.

Djuke Veldhuis. 8 scientific storytelling. *The Aesthetics of Scientific Data Representation: More than Pretty Pictures*, page 31, 2017.

A. Voinov, N. Kolagani, M.K. McCall, P.D. Glynn, M.E. Kragt, F.O. Ostermann, S.A. Pierce, and P. Ramu. Modelling with stakeholders - next generation. *Environmental Modelling and Software*, 77:196–220, 2016.

X. Wang, L. Tian, B. Xu, X. Wang, and W. Wu. Mooc for medical big data research: An important role in hypertension big data research. In *2015 IEEE FIRST INTERNATIONAL CONFERENCE ON BIG DATA COMPUTING SERVICE AND APPLICATIONS (BIGDATASERVICE 2015)*, pages 453–455. IEEE; IEEE Comp Soc; San Jose Univ; Beihang Univ; ITC NSTI; Arizona State Univ; Peking Univ; NE Polytechn Univ; Deakin Univ Australia; UOttawa; UTS; Univ Leeds; Taiyuan Univ Technol; Tsinghua Univ, 2015. ISBN 978-1-4799-8128-1. 1st IEEE International Conference on Big Data Computing Service and Applications (BigDataService), San Francisco, CA, MAR 30-APR 03, 2015.

LEON WATTS and TOM WRIGGLESWORTH. Emotional connections with the past: exploring engagement with historical images from an online museum collection. In *Cultural Heritage Communities*, pages 152–169. Routledge, 2017.

S. Wiemann, P. Karrasch, and L. Bernard. Ad-hoc combination and analysis of heterogeneous and distributed spatial data for environmental monitoring –design and prototype of a web-based solution. *International Journal of Digital Earth*, 11(1):79–94, 2018. ISSN 1753-8947.

Yolanda F Wiersma. What kind of scientist are you? science and interdisciplinary research. *Scholarly and Research Communication*, 5(1), 2013.

Jodi Christine Wilker. Women in stem: The effect of undergraduate research on persistence. 2017.

Julia Wilkinson, Chris J Scott, and David M Willis. Going with the floe. *Astronomy & Geophysics*, 57(2):2–37, 2016.

Marcus Winter. Topical objects study. 2014.

Walter Witschey, Howell J Parry Jr, Eugene Maurakis, David Hagan, Maia Werner-Avidon, Chuck Howarth, and Don Pohlman. Museums in transition: emerging technologies as tools for free-choice learning. *Informal Learning Review, The*, (81), 2006.

Tom Wrigglesworth and Leon Watts. Exploring engagement with historical images from an online museum collection. *Cultural Heritage Communities: Technologies and Challenges*, 2017.

Li Zhao, Brian Detlor, and Catherine E Connelly. Sharing knowledge in social q&a sites: the unintended consequences of extrinsic motivation. *Journal of Management Information Systems*, 33(1): 70–100, 2016.

Y. Zhou and Y. Long. Sinogrids: a practice for open urban data in china. *Cartography and Geographic Information Science*, 43(5):379–392, November 2016. ISSN 1523-0406.

Yun Zhu, Leisi Pei, and Junjie Shang. Improving video engagement by gamification: A proposed design of mooc videos. In *International Conference on Blended Learning*, pages 433–444. Springer, 2017.

A.4 Methodologies

These papers concerned methodologies for Citizen Science projects, or for the use of Citizen Science data:

Mark Andersen and Gary Beauvais. Predictive distribution modeling of species of greatest conservation need in texas. *Report prepared by the Wyoming Natural Diversity Database., University of Wyoming, Laramie, WY*, 2013.

Victor Anton, Stephen Hartley, Andre Geldenhuis, and Heiko U Wittmer. Monitoring the mammalian fauna of urban areas using remote cameras and citizen science. *Journal of Urban Ecology*, 4(1): juy002, 2018.

J. Arsanjania and E. Vaz. An assessment of a collaborative mapping approach for exploring land use patterns for several european metropolises. *International Journal of Applied Earth Observation and Geoinformation*, 35(PB): 329–337, March 2015. ISSN 0303-2434.

M. BÃl, J. KubeÄek, J. SedonÃk, and R. AndrÃÅik. Srazenazver.cz: A system for evidence of animal-vehicle collisions along transportation networks. *Biological Conservation*, 213(A):167–174, September 2017. ISSN 0006-3207.

R.J. Beeden, M.A. Turner, J. Dryden, F. Merida, K. Goudkamp, C. Malone, P.A. Marshall, A. Birtles, and J.A. Maynard. Rapid survey protocol that provides dynamic information on reef condition to managers of the great barrier reef. *Environmental Monitoring and Assessment*, 186(12):8527–8540, December 2014. ISSN 0167-6369.

P. Bennett-Martin, C.C. Visaggi, and T.L. Hawthorne. Mapping marine debris across coastal communities in belize: developing a baseline for understanding the distribution of litter on beaches using geographic information systems. *Environmental Monitoring and Assessment*, 188(10), October 2015. ISSN 0167-6369.

Marguerite Benony, Marianne Cardon, Arnaud Ferré, Jean Coquet, Nathan Foulquier, Florian Thonier, Lucas Le Lann, Henry De Belly, Alexandre Evans, Aakriti Jain, et al. The smell of us–crowdsourcing human body odor evaluation. *Human Computation*, 3(1): 161–179, 2016.

Eskender Beza, Eliakim Hamunyela, Pytrik Reidsma, Melisew Misker, and Lammert Kooistra Belay. Remote sensing and crowdsourcing for estimating and explaining yields: sesame in ethiopia. *Citizen science and remote sensing for crop yield gap analysis*, page 119.

Glenn A Bowen. Document analysis as a qualitative research method. *Qualitative research journal*, 9(2):27–40, 2009.

Connor Bowley, Marshall Mattingly, Andrew Barnas, Susan Ellis-Felege, and Travis Desell. Toward using citizen scientists to drive automated ecological object detection in aerial imagery. In *e-Science (e-Science), 2017 IEEE 13th International Conference on*, pages 99–108. IEEE, 2017.

G Bracey. Evaluation of a virtual citizen science facility: A comprehensive mixed-methods approach. In JG Manning, MK Hemenway, JB Jensen, and MG Gibbs, editors, *Ensuring Stem Literacy: A National Conference on STEM Education and Public Outreach*, volume 483 of *Astronomical Society of the Pacific Conference Series*, page 233. EXOPLANET EXPLORAT PROGRAM; JPL; Stratospher Observ Infrared Astron; NRAO; E&S; Sky Skan; LOCKHEED MARTIN; sapl learn; Bell Aerosp & Technol Corp; AAS; EXPLORE SCI; SEILER INSTRUMENT; Astron Soc Pacif; San Jose State Univ, 2014. ISBN 978-1-58381-850-3. 125th ASP Annual Conference on Ensuring STEM Literacy: A National Conference on STEM Education and Public Outreach, San Jose State Univ, San Jose, CA, JUL 20-24, 2013.

CHRISTOPHER J BRADY. Crowdsourcing the analysis of retinal imaging data.

Christopher J Brady. Crowdsourcing for rapid fundus photograph interpretation. *Retina Today*, (APRIL 2015):71–73, 2015.

Christopher J Brady, Andrea C Villanti, Jennifer L Pearson, Thomas R Kirchner, Omesh P Gupta, and Chirag P Shah. Rapid grading of fundus photographs for diabetic retinopathy using crowdsourcing. *Journal of medical Internet research*, 16(10), 2014.

Tim B. Brown, Kevin R. Hultine, Heidi Steltzer, Ellen G. Denny, Michael W. Denslow, Joel Granados, Sandra Henderson, David Moore, Shin Nagai, Michael SanClements, Arturo Sanchez-Azofeifa, Oliver Sonnentag, David Tazik, and Andrew D. Richardson. Using phenocams to monitor our changing earth: toward a global phenocam network. *FRONTIERS IN ECOLOGY AND THE ENVIRONMENT*, 14(2):84–93, March 2016. ISSN 1540-9295.

Jarrett Byrnes, KC Cavanaugh, AJ Haupt, L Trouille, I Rosenthal, TW Bell, A Rassweiler, A Pérez-Matus, and J Assis. Using online citizen science to assess giant kelp abundances across the globe with satellite imagery. In *AGU Fall Meeting Abstracts*, 2017.

N.A. Case, E.A. MacDonald, and R. Viereck. Using citizen science reports to define the equatorial extent of auroral visibility. *Space Weather*, 14(3): 198–209, March 2016. ISSN 1542-7390.

Kevin RV Casteels, Steven P Bamford, Ramin A Skibba, Karen L Masters, Chris J Lintott, William C Keel, Kevin Schawinski, Robert C Nichol, and Arfon M Smith. Galaxy zoo: quantifying morphological indicators of galaxy interaction. *Monthly Notices of the Royal Astronomical Society*, 429(2):1051–1065, 2012.

Steven D Collins, John C Abbott, and Nancy E McIntyre. Quantifying the degree of bias from using county-scale data in species distribution modeling: Can increasing sample size or using county-averaged environmental data reduce distributional overprediction? *Ecology and evolution*, 7(15):6012–6022, 2017.

T.K. Davies, G. Stevens, M.G. Meekan, J. Struve, and J.M. Rowcliffe. Can citizen science monitor whale-shark aggregations? investigating bias in mark-recapture modelling using identification photographs sourced from the public. *Wildlife Research*, 39(8):696–704, 2012. ISSN 1035-3712.

Darren R. Davis and Wayne B. Hayes. Automated quantitative description of spiral galaxy arm-segment structure. In *2012 IEEE CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION (CVPR)*, IEEE Conference on Computer Vision and Pattern Recognition, pages 1138–1145. IEEE, 2012. ISBN 978-1-4673-1228-8. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Providence, RI, JUN 16-21, 2012.

K. Dehnen-Schmutz and J. Conroy. Working with gardeners to identify potential invasive ornamental garden plants: testing a citizen science approach. *Biological Invasions*, pages 1–9, 2018.

A.J. Dennhardt, A.E. Duerr, D. Brandes, and T.E. Katzner. Applying citizen-science data and mark-recapture models to estimate numbers of migrant golden eagles in an important bird area in eastern north america. *Condor*, 119 (4):817–831, November 2017. ISSN 0010-5422.

Mark Desholm, Rashpal Gill, Thomas Bøvith, and Anthony D Fox. Combining spatial modelling and radar to identify and protect avian migratory hot-spots. *Current Zoology*, 60(5):680–691, 2014.

Julian R. Dupuis, Christianne M. Mcdonald, John H. Acorn, and Felix A. H. Sperling. Genomics-informed species delimitation to support morphological identification of anglewing butterflies (lepidoptera: Nymphalidae: Polygonia). *ZOOLOGICAL JOURNAL OF THE LINNEAN SOCIETY*, 183(2):372–389, June 2018. ISSN 0024-4082.

C. Ellul, L. Francis, and M. Haklay. A flexible database-centric platform for citizen science data capture. pages 39–44, 2011.

Daniel Fink, Wesley M. Hochachka, Benjamin Zuckerberg, David W. Winkler, Ben Shaby, M. Arthur Munson, Giles Hooker, Mirek Riedewald, Daniel Sheldon, and Steve Kelling. Spatiotemporal exploratory models for broad-scale survey data. *Ecological Applications*, 20 (8):2131–2147, December 2010. ISSN 10510761.

William Fithian, Jane Elith, Trevor Hastie, and David A Keith. Bias correction in species distribution models: pooling survey and collection data for multiple species. *Methods in Ecology and Evolution*, 6(4):424–438, 2015.

E.A. Forys and A.R. Hevesh. Investigating black skimmer chick diets using citizen science and digital photography. *Southeastern Naturalist*, 16(3):317–325, September 2017. ISSN 1528-7092.

N.R. Foster, D.G. Fotheringham, D.J. Brock, and M. Waycott. A resourceful and adaptable method to obtain data on the status of seagrass meadows. *Aquatic Botany*, 141:17–21, July 2017. ISSN 0304-3770.

A. Friedrichs, J.A. Busch, H.J. van der Woerd, and O. Zielinski. Smartfluo: A method and affordable adapter to measure chlorophyll a fluorescence with smartphones. *Sensors (Switzerland)*, 17 (4), April 2017. ISSN 1424-8220.

Gurutzeta Guillera-Arroita, José J Lahoz-Monfort, Jane Elith, Ascelin Gordon, Heini Kujala, Pia E Lentini, Michael A McCarthy, Reid Tingley, and Brendan A Wintle. Is my species distribution model fit for purpose? matching data and models to applications. *Global Ecology and Biogeography*, 24(3):276–292, 2015.

Courtney Hann. Evaluation of the marine debris tracker app on the us west coast.

Heather J Henter, Ralph Imondi, Karen James, Diana Spencer, and Dirk Steinke. Dna barcoding in diverse educational settings: five case studies. *Phil. Trans. R. Soc. B*, 371(1702):20150340, sep 5 2016. ISSN 0962-8436.

H. Hutt, R. Everson, M. Grant, J. Love, and G. Littlejohn. How clumpy is my image?: Scoring in crowdsourced annotation tasks. *Soft Computing*, 19(6): 1541–1552, 2015.

G.V. Ionescu, E.F. Harkness, J. Hulleman, and S.M. Astley. A citizen science approach to optimising computer aided detection (cad) in mammography. volume 10577, 2018.

Susan Jamieson. Likert scales: how to (ab) use them. *Medical education*, 38(12): 1217–1218, 2004.

C.M. Johnson, L.E. Beckley, H. Kobryn, G.E. Johnson, I. Kerr, and R. Payne. Crowdsourcing modern and historical data identifies sperm whale (physeter macrocephalus) habitat offshore of south-western australia. *Frontiers in Marine Science*, 3(SEP), 2016.

A.M. Jordt, M. Lange, S. Kramer-Schadt, L.H. Nielsen, S.S. Nielsen, H.-H. Thulke, H. Vejre, and L. Alban. Spatio-temporal modeling of the invasive potential of wild boar-a conflict-prone species-using multi-source citizen science data. *Preventive Veterinary Medicine*, 124:34–44, feb 1 2016. ISSN 0167-5877.

J. Kamp, S. Oppel, H. Heldbjerg, T. Nyegaard, and P.F. Donald. Unstructured citizen science data fail to detect long-term population declines of common birds in denmark. *Diversity and Distributions*, 22(10):1024–1035, October 2016. ISSN 1366-9516.

N. Kramer and E. Wohl. Rules of the road: A qualitative and quantitative synthesis of large wood transport through drainage networks. *Geomorphology*, 279(SI):74–97, feb 15 2017. ISSN 0169-555X. 3rd International Conference on Wood in World Rivers, Padova, ITALY, JUL, 2015.

J. Mas, A. Fornes, and J. Llados. An interactive transcription system of census records using word-spotting based information transfer. pages 54–59, 2016.

Robert M McElderry. Estimating density and temperature dependence of juvenile vital rates using a hidden markov model. *Insects*, 8(2):51, 2017.

E.K. Melaas, M.A. Friedl, and A.D. Richardson. Multiscale modeling of spring phenology across deciduous forests in the eastern united states. *Global Change Biology*, 22(2):792–805, February 2016. ISSN 1354-1013.

David A. W. Miller, Linda A. Weir, Brett T. McClintock, Evan H. Campbell Grant, Larissa L. Bailey, and Theodore R. Simons. Experimental investigation of false positive errors in auditory species occurrence surveys. *ECOLOGICAL APPLICATIONS*, 22(5): 1665–1674, July 2012. ISSN 1051-0761.

E Minkman, MM Rutten, and MCA van der Sanden. Aq methodological approach to identify practitioners' viewpoints on citizen science in dutch regional water resource management. *Hydrology and Earth System Sciences*, 20 (1):1–1, 2016.

G. Nelson, P. Sweeney, L.E. Wallace, R.K. Rabeler, D. Allard, H. Brown, J.R. Carter, M.W. Denslow, E.R. Ellwood, C.C. Germain-Aubrey, E. Gilbert, E. Gillespie, L.R. Goertzen, B. Legler, D.B. Marchant, T.D. Marsico, A.B. Morris, Z. Murrell, M. Nazaire, C. Neefus, S. Oberreiter, D. Paul, B.R. Ruhfel, T. Sasek, J. Shaw, P.S. Soltis, K. Watson, A. Weeks, and A.R. Mast. Digitization workflows for flat sheets and packets of plants, algae, and fungi. *Applications in Plant Sciences*, 3(9), September 2015. ISSN 2168-0450.

X.J. Nelson and N. Fijn. The use of visual media as a tool for investigating animal behaviour. *Animal Behaviour*, 85(3):525–536, March 2013. ISSN 0003-3472.

Senta Niederegger, Klaus-Peter Döge, Marcus Peter, Tobias Eickhölter, and Gita Mall. Connecting the dots: From an easy method to computerized species determination. *Insects*, 8(2):52, 2017.

T. Osawa, T. Yamanaka, Y. Nakatani, J. Nishihiro, S. Takahashi, S. Mahoro, and H. Sasaki. A crowdsourcing approach to collecting photobased insect and plant observation records. *Biodiversity Data Journal*, 5, 2017.

J. Pagel, B.J. Anderson, R.B. O'Hara, W. Cramer, R. Fox, F. Jeltsch, D.B. Roy, C.D. Thomas, and F.M. Schurr. Quantifying range-wide variation in population trends from local abundance surveys and widespread opportunistic occurrence records. *Methods in Ecology and Evolution*, 5(8):751–760, August 2014. ISSN 2041-210X.

M. Peltoniemi, M. Aurela, K. BÃ¶ttcher, P. Kolari, J. Loehr, J. Karhu, M. Linkosalmi, C.M. Tanis, J.-P. Tuovinen, and A.N. Arslan. Webcam network and image database for studies of phenological changes of vegetation and snow cover in finland, image time series from 2014 to 2016. *Earth System Science Data*, 10(1):173–184, jan 25 2018. ISSN 1866-3508.

A. Pilny, B. Keegan, B.F. Wells, C. Riedl, D. Lazer, J. Radford, K. Ognyanova, L. DeChurch, M. Macy, N. Contractor, and W. Meleis. Designing online experiments: Citizen science approaches to research. volume 26-February-2016, pages 498–502, 2016.

J.D. Ramsdale, M.R. Balme, S.J. Conway, C. Gallagher, S.A. van Gasselt, E. Hauber, C. Orgel, A. SÃjourn A, J.A. Skinner, F. Costard, A. Johnsson, A. Losiak, D. Reiss, Z.M. Swirad, A. Kereszturi, I.B. Smith, and T. Platz. Grid-based mapping: A method for rapidly determining the spatial distributions of small features over very large areas. *Planetary and Space Science*, 140: 49–61, June 2017. ISSN 0032-0633.

Zuriel Anne Rasmussen. Coyotes on the web: Understanding human-coyote interaction and online education using citizen science. 2015.

Jonathan P Rose and Brian D Todd. Projecting invasion risk of non-native watersnakes (nerodia fasciata and nerodia sipedon) in the western united states. *PLoS One*, 9(6):e100277, 2014.

Helen E Roy, Elizabeth Baxter, Aoine Saunders, and Michael JO Pocock. Focal plant observations as a standardised method for pollinator monitoring: Opportunities and limitations for mass participation citizen science. *PloS one*, 11(3):e0150794, 2016.

J. Rudd, N. Stephenson, D.L. Clifford, L. Konde, J.T. Villepique, and J. Foley. Utilizing citizen science to document a mange epidemic in western gray squirrels in california. *Wildlife Society Bulletin*, 40(2):261–268, June 2016. ISSN 1938-5463.

Alison M. Smith and Paul M. Ramsay. A comparison of ground-based methods for estimating canopy closure for use in phenology research. *AGRICULTURAL AND FOREST METEOROLOGY*, 252: 18–26, apr 15 2018. ISSN 0168-1923.

Patrícia Tiago, Henrique M Pereira, and César Capinha. Using citizen science data to estimate climatic niches and species distributions. *Basic and Applied Ecology*, 20:75–85, May 2017. ISSN 1439-1791.

Brian D Todd, A Justin Nowakowski, Jonathan P Rose, and Steven J Price. Species traits explaining sensitivity of snakes to human land use estimated from citizen science data. *Biological conservation*, 206:31–36, 2017.

Brian D Todd, Jonathan P Rose, Steven J Price, and Michael E Dorcas. Using citizen science data to identify the sensitivity of species to human land use. *Conservation biology*, 30(6): 1266–1276, 2016.

Jacob van Etten. Crowdsourcing crop improvement in sub-saharan africa: a proposal for a scalable and inclusive approach to food security. *IDS bulletin*, 42 (4):102–110, July 2011. ISSN 0265-5012.

Jacob Van Etten, Eskender Beza, Lluís Calderer, Kees Van Duijvendijk, Carlo Fadda, Basazen Fantahun, Yosef Gebrehawaryat Kidane, Jeske van de Gevel, Arnab Gupta, Dejene Kassahun Mengistu, et al. First experiences with a novel farmer citizen science approach: crowdsourcing participatory variety selection through on-farm triadic comparisons of technologies (tricot). *Experimental Agriculture*, pages 1–22, 2016.

Arco J van Strien, Chris AM van Swaay, and Tim Termaat. Opportunistic citizen science data of animal species produce reliable estimates of distribution trends if analysed with occupancy models. *Journal of Applied Ecology*, 50 (6):1450–1458, 2013.

Matt von Konrat, Thomas Campbell, Ben Carter, Matthew Greif, Mike Bryson, Juan Larraín, Laura Trouille, Steve Cohen, Eve Gaus, Ayesha Qazi, et al. Using citizen science to bridge taxonomic discovery with education and outreach. *Applications in plant sciences*, 6(2):e1023, February 2018. ISSN 2168-0450.

M Wesley, W David, M Arthur, et al. Spatiotemporal exploratory models for broadâ scale survey data. *Ecological applications*, 2010.

A.C. Williams, J.F. Wallin, H. Yu, M. Perale, H.D. Carroll, A.-F. Lamblin, L. Fortson, D. Obbink, C.J. Lintott, and J.H. Brusuelas. A computational pipeline for crowdsourced transcriptions of ancient greek papyrus fragments. In

J Lin, XH Hu, W Chang, R Nambiar, C Aggarwal, N Cercone, V Honavar, J Huan, B Mobasher, and S Pyne, editors, *2014 IEEE INTERNATIONAL CONFERENCE ON BIG DATA (BIG DATA)*, pages 100–105. IEEE; IEEE Comp Soc; ELSEVIER; Natl Sci Fdn; CISCO; CCF, 2015. ISBN 978-1-4799-5666-1. IEEE International Conference on Big Data, Washington, DC, OCT 27-30, 2014.

A.T. Woods, C. Velasco, C.A. Levitan, X. Wan, and C. Spence. Conducting perception research over the internet: A tutorial review. *PeerJ*, 2015(7), jul 23 2015. ISSN 2167-8359.

Jennifer M Yost, Patrick W Sweeney, Ed Gilbert, Gil Nelson, Robert Guralnick, Amanda S Gallinat, Elizabeth R Ellwood, Natalie Rossington, Charles G Willis, Stanley D Blum, et al. Digitization protocol for scoring reproductive phenology from herbarium specimens of seed plants. *Applications in plant sciences*, 6(2):e1022, 2018.

M. Zubairu. Community remote sensing for maximum spatial data interaction. volume 4, pages 3229–3232, 2012.

A.5 My Publications

Peer-Reviewed Publications in which I was involved as an author:

Neal Reeves, Ramine Tinati, Sergej Zerr, Elena Simperl, and Max Van Kleek. From crowd to community: a survey of online community features in citizen science projects. *Proceedings of the ACM Conference on Computer Supported Cooperative Work, CSCW*, pages 2137–2152, 2017.

Neal Reeves, Peter West, and Elena Simperl. "a game without competition is hardly a game": The impact of competitions on player activity in a human computation game. AAAI, 2018.

Elena Simperl, Neal Reeves, Chris Phethean, Todd Lynes, and Ramine Tinati. Is virtual citizen science a game? *ACM Transactions on Social Computing*, 1(2):6, 2018.

A.6 Not Peer Reviewed

Research articles which are not peer reviewed (e.g., PhD/Masters thesis):

Morgan Adams. *Evaluating the role of citizen science in the context of human-wildlife conflict management.* PhD thesis, Colorado State University. Libraries, 2014.

Stuart Anderson. inaturalist: Understanding biodiversity through a digital medium. Master's thesis, University of Waterloo, 2018.

Maria Aristeidou. *Citizen Inquiry: Engaging citizens in online communities of scientific inquiries.* PhD thesis, Institute of Educational Technology, 2016.

ME Balestrini. *A City in Common: Explorations on Sustained Community Engagement with Bottom-up Civic Technologies.* PhD thesis, UCL (University College London), 2017.

David Becerra. *Predicting Protein Folding Pathways Using Ensemble Modeling and Sequence Information.* PhD thesis, McGill University, 2017.

Eskender Andualem Beza. *Citizen science and remote sensing for crop yield gap analysis.* PhD thesis, Wageningen University, 2017.

Anne Elizabeth Bowser. *Cooperative design, cooperative science: Investigating collaborative research through design with floracaching.* PhD thesis, 2015.

Kevin Casteels. *The Evolution, Masses and Morphologies of Merging Galaxies.* PhD thesis, Universitat de Barcelona, 2012.

Sarah K Chase. *Evaluating citizen science for natural resource monitoring: Participant motivations, attitude and behavioral change, and the influence of natural resource characteristics on program potential.* PhD thesis, San Diego State University, 2016.

Steven D Collins. *Fine-scale modeling of riverine Odonata distributions in the northeastern United States.* PhD thesis, 2014.

DITOs Consortium et al. Doing it together science: D6. 2 initial plan for communications, dissemination and exploitation. 2016.

DITOs Consortium et al. Doing it together science: D6. 5 plan for communications, dissemination and exploitation-update. 2017.

Heather Heather Hult Craig. *Interactive data narrative: Designing for public engagement.* PhD thesis, Massachusetts Institute of Technology, 2015.

Martin Sebastian Dittus. *Analysing Volunteer Engagement in Humanitarian Crowdmapping.* PhD thesis, UCL (University College London), 2017.

Timothy Dobosz. *Increasing Citizen Science Participation in the Picture Post Project.* PhD thesis, WORCESTER POLYTECHNIC INSTITUTE, 2011.

João André Ribeiro Duarte. *Can we play science?: philosophical perspectives on participation in science research.* PhD thesis, 2015.

Anne Edwards. *Examining the structures and practices for knowledge production within Galaxy Zoo-an online citizen science initiative.* PhD thesis, University of Oxford, 2014.

Alexandra Margaret Mary Eveleigh. *Crowding out the Archivist? Implications of online user participation for archival theory and practice.* PhD thesis, UCL (University College London), 2015.

Whitney Ferrin-Rodriguez. *The hybrid laboratory: Informal spaces for public-science interaction.* PhD thesis, Loyola University Chicago, 2012.

David Thomas Tyler Flockhart. *Population Dynamics of a Long-distance Migratory Insect.* PhD thesis, 2013.

Hoda Hamouda. *Eyewitness: Platform Design for Visualizing and Synthesizing Citizen Media Video Content of Political Significance.* PhD thesis, UNIVERSITY OF ART, 2014.

Teresa Holocher-Ertl and Barbara Kieslinger ZSI. Deliverable no. d5. 3 deliverable name draft white paper (green paper) dissemination level pu wp no. 5 wp name evaluation and policy recommendations. 2013.

Jo Iacovides, Charlene Jennett, Cassandra Cornish-Trestrail, Anna Cox, I Iacovides, C Jennett, C Cornish-Trestrail, and AL Cox. Public paper no. 111.

Rebecca M Jarvis. *Putting people back in the picture: a social research agenda for a social-ecological approach to conservation planning.* PhD thesis, Auckland University of Technology, 2015.

Andrew James Blevins Jennings. *The effects of preformed scour holes on anuran biodiversity in the North Central Piedmont region.* PhD thesis, The University of North Carolina at Greensboro, 2013.

Larry Johnson, S Adams Becker, Victoria Estrada, Alex Freeman, Panagiotis Kampylis, Riina Vuorikari, and Yves Punie. Nmc horizon report europe: 2014 schools edition. Technical report, The New Media Consortium, 2014.

Sun Young Kim. *Democratizing mo-*

bile technology in support of volunteer activities in data collection. PhD thesis, Human-Computer Interaction Institute, 2010.

Jana Lee Kubecka. *The Influence of Knowledge Gained and the Likelihood of Recommending Texas A&M Agrilife Extension Service on the Planned Adoption of Wild Pig Control Techniques.* PhD thesis, 2016.

E. Law, A.C. Williams, J. Shirk, A. Wiggins, J. Brier, J. Preece, and G. Newman. The science of citizen science: Theories, methodologies and platforms. pages 395–400, 2017.

Roman Lukyanenko. Best dissertation in is in 2015: An information modeling approach to improve quality of user-generated content.

Louise I Lynch. *Science experiences of citizen scientists in entomology research.* PhD thesis, The University of Nebraska-Lincoln, 2016.

ET Meyer. *e-Research in the life sciences: from invisible to virtual colleges.* PhD thesis, University of Oxford, 2011.

Lauren Neville. *Rhetoric of Beekeeping: Science, Technology & Innovation.* PhD thesis, Georgetown University, 2017.

Miguel Querejeta. *Making galaxies passive: Insights from resolved studies of nearby galaxies.* PhD thesis, 2016.

Sriswetha Rajagopal. *A Human Computation Approach to Graph Isomorphism.* PhD thesis, McGill University Libraries, 2016.

Mia Ridge. *Making digital history: The impact of digitality on public participation and scholarly practices in historical research.* PhD thesis, Open University, 2016.

Courtney Rieger et al. *Demonstrating the capacity of online citizen science mapping software to communicate natural hazards and engage community participation.* PhD thesis, Lethbridge, Alta: University of Lethbridge, Dept. of Geography, 2016.

Bruno Padilha Rocha, Flávio Esplugues Sanches Calegari, and Jordan Kaique Kobellarz. Fatores e mecanismos motivadores em comunidades virtuais. B.S. thesis, Universidade Tecnológica Federal do Paraná, 2017.

Calum Sharp et al. *Human response to environmental noise and vibration from freight and passenger railway traffic.* PhD thesis, University of Salford, 2014.

Vidya Bhushan Singh. *User modeling and optimization for environmental planning system design.* PhD thesis, 2014.

A.M. Smith, S. Lynn, and C.J. Lintott. An introduction to the zooniverse. volume WS-13-18, page 103, 2013.

James C Sprinks. *Designing task workflows to ensure the best scientific outcomes in citizen science.* PhD thesis, University of Nottingham, 2017.

Jonathan Steinke. Citizen science with resource-poor farmers as a new approach to climate adaption and food security: Evidence from honduras. Master's thesis, Humboldt-Universität zu Berlin, Lebenswissenschaftliche Fakultät, 2015.

Kristine F Stepenuck. *Improving understanding of outcomes and credibility of volunteer environmental monitoring programs.* PhD thesis, UNIVERSITY OF WISCONSIN-MADISON, 1913.

Quasar Surprise. *Student as Scientist: Measuring Outcomes of Contributory and Collaborative Citizen Science.* PhD thesis, Evergreen State College, 2017.

Sean R Tracy. *Inclusion of Environmental Education into Public School Curricula.* PhD thesis, George Mason University, 2017.

SC van der Kraan. Amsterdamse monumenten. de verhouding tussen historicus en leek in een digitale participatory historical culture. Master's thesis, 2016.

W.-P. Vellinga and R. PlanquÃ. The xeno-canto collection and its relation to sound recognition and classification. volume 1391, 2015.

Katherine Bibee Wolfson. *Citizen science: A tool for scientific discovery and influencing science attitudes in museums.* PhD thesis, University of Colorado at Boulder, 2015.

Reviewers Saskia Woutersen-Windhouwer and Eleni Toli. Deliverable d4. 1–practices evaluation and mapping: Meth-ods, tools and user needs.

Charmian L Zoll. Leveraging transmedia communication strategies to improve engagement and foster collaboration in citizen-science projects. B.S. thesis, University of Waterloo, 2017.

A.7 Not VCS

Publications which describe offline forms of Citizen Science:

John H Acorn. Entomological citizen science in canada. *The Canadian Entomologist*, 149(6):774–785, December 2017. ISSN 0008-347X.

Bethany J Alender. *Understanding volunteer motivations to participate in citizen science projects: a deeper look at water quality monitoring.* PhD thesis, Evergreen State College, 2015.

Stuart Allan and Jacqui Ewart. New forms of environmental reporting. *The Routledge Handbook of Environment and Communication*, page 186, 2015.

Paul Aoki, Allison Woodruff, Baladitya Yellapragada, and Wesley Willett. Environmental protection and agency: Motivations, capacity, and goals in participatory sensing. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pages 3138–3150. ACM, 2017.

JC AUBELE. Citizen science in a museum setting. In *2010 GSA Denver Annual Meeting*, 2010.

R. Bardaji, M. Best, and J. Piera. Citizen science in high-latitude ecosystems: Kduino buoy to support data gathering in extreme environments. *Sea Technology*, 58(2):37–39, 2017.

A. Baumber, G. Metternicht, P. Ampt, R. Cross, and E. Berry. Opportunities for adaptive online collaboration to enhance rural land management. *Journal of Environmental Management*, 219: 28–36, 2018.

Eskender Beza, Pytrik Reidsma, P Marijn Poortvliet, Melisew Misker Belay, Ben Sjors Bijen, and Lammert Kooistra. Exploring farmers' intentions to adopt mobile short message service (sms) for citizen science in agriculture. *Computers and Electronics in Agriculture*, 151: 295–310, 2018.

Eskender Beza, Jonathan Steinke, Jacob Van Etten, Pytrik Reidsma, Carlo Fadda, Sarika Mittra, Prem Mathur, and Lammert Kooistra. What are the prospects for citizen science in agriculture? evidence from three continents on motivation and mobile telephone use of resource-poor farmers. *PloS one*, 12(5): e0175700, 2017.

Susanne Bødker, Henrik Korsgaard, and Joanna Saad-Sulonen. 'a farmer, a place and at least 20 members': The development of artifact ecologies in volunteer-based communities. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*, pages 1142–1156. ACM, 2016.

J.R. Brammer, N.D. Brunet, A.C. Burton, A. Cuerrier, F. Danielsen, K. Dewan, T.M. Herrmann, M.V. Jackson, R. Kennett, G. Larocque, M. Mulrennan, A.K. Pratihast, M. Saint-Arnaud, C. Scott, and M.M. Humphries. The role of digital data entry in participatory environmental monitoring. *Conservation Biology*, 30(6):1277–1287, December 2016. ISSN 0888-8892.

Robert JW Brewin, Lee de Mora, Thomas Jackson, Thomas G Brewin, and Jamie Shutler. On the potential of surfers to monitor environmental indicators in the coastal zone. *PloS one*, 10 (7):e0127706, 2015.

J.A. Busch, R. Bardaji, L. Ceccaroni, A. Friedrichs, J. Piera, C. Simon, P. Thijsse, M. Wernand, H.J. van derWoerd, and O. Zielinski. Citizen bio-optical observations from coast- and ocean and their compatibility with ocean colour satellite measurements. *Remote Sensing*, 8(11), 2016.

C. Camacho. Birding trip reports as a data source for monitoring rare species. *Animal Conservation*, 19(5):430–435, October 2016. ISSN 1367-9430.

Jessica L Cappadonna, Margot Brereton, David M Watson, and Paul Roe. Calls from the wild: Engaging citizen scientist with animal sounds. In *Proceedings of the 2016 ACM Conference Companion Publication on Designing Interactive Systems*, pages 157–160. ACM, 2016.

Carlos Carcia-Soto and Gro Ingleid

van der Meeren. Advancing citizen science for coastal and ocean research. *European Marine Board IVZW*, 2017.

Chandara Chea et al. Gamified participatory sensing: Impact of gamification on public's motivation in a lake monitoring application. 2017.

Roxana Ciceoi, Liliana-Aurelia Bădulescu, Minodora Gutue, Elena Ştefania Mardare, Cristian Mihai Pomohaci, et al. Citizen-generated data on invasive alien species in romania: Trends and challenges. *Acta Zoologica Bulgarica, Supplement*, 9:255–260, 2017.

SURFACING INTUITION THROUGH A CITIZEN. Chapter four making knowledge and surfacing intuition through a citizen science game ladan cockshut. *Landscapes of Participatory Making, Modding and Hacking: Maker Culture and Makerspaces*, page 55, 2017.

AL Conway, GT Green, L Larson, SM Hernandez, and Carroll JP. A new context for citizen science?: Using local knowledge to inform wildlife management and conservation in sierra leone, africa. *APRIL LEANNE CONWAY*, page 55.

Caren Cooper, Lincoln Larson, Kathleen Krafte Holland, Rebecca Gibson, David Farnham, Diana Hsueh, Patricia Culligan, and Wade McGillis. Contrasting the views and actions of data collectors and data consumers in a volunteer water quality monitoring project: Implications for project design and management. *Citizen Science: Theory and Practice*, 2(1), 2017.

Laura E Coristine, Peter Soroye, Rosana Nobre Soares, Cassandra Robillard, and Jeremy T Kerr. Dispersal limitation, climate change, and practical tools for butterfly conservation in intensively used landscapes. *Natural areas journal*, 36(4):440–452, 2016.

Jason R Courter, Ron J Johnson, Claire M Stuyck, Brian A Lang, and Evan W Kaiser. Weekend bias in citizen science data reporting: implications for phenology studies. *International journal of biometeorology*, 57(5):715–720, 2013.

Rhiannon Crain, Caren Cooper, and Janis L Dickinson. Citizen science: a tool for integrating studies of human and natural systems. *Annual Review of Environment and Resources*, 39, 2014.

Davi GF Cunha, Jonatas F Marques, JULIANA C RESENDE, PATRÍCIA B FALCO, CHRISLAINE M SOUZA, and Steven A Loiselle. Citizen science participation in research in the environmental sciences: key factors related to projects' success and longevity. *Anais da Academia Brasileira de Ciências*, 89(3): 2229–2245, 2017.

K. Dehnen-Schmutz, G.L. Foster, L. Owen, and S. Persello. Exploring the role of smartphone technology for citizen science in agriculture. *Agronomy for Sustainable Development*, 36(2), June 2016. ISSN 1774-0746.

J. Del Rio, J. Aguzzi, A. Hidalgo, I. Bghiel, A. Manuel, V. Sbragaglia, and F. Sarda. Citizen science and marine community monitoring by video-cabled observatories: The obsea citizen science project. In *2013 IEEE INTERNATIONAL UNDERWATER TECHNOLOGY SYMPOSIUM (UT)*. IEEE, 2013. ISBN 978-1-4673-5948-1; 978-1-4673-5947-4. IEEE International Underwater Technology Symposium (UT), Tokyo, JAPAN, MAR 05-08, 2013.

Erik DeLuca. Wolf listeners: An introduction to the acoustemological politics and poetics of isle royale national park. *Leonardo Music Journal*, pages 87–90, 2016.

C. Deutsch, D. Bilenca, and G. Agostini. In search of the horned frog (ceratophrys ornata) in argentina: Complementing field surveys with citizen science. *Herpetological Conservation and Biology*, 12(3):664–672, December 2017. ISSN 2151-0733.

William G. Deutsch and Sergio S. Ruiz-C A^{3}rdova. Trends, challenges, and responses of a 20-year, volunteer water monitoring program in alabama. *Ecology and Society*, 20(3), 2015. ISSN 17083087.

A. Dolan. Citizen science and smart phone emissions monitoring. pages 318–322, 2016.

S. Dolan. Awma ace citizen science emission monitoring. 2017.

Margret C Domroese and Elizabeth A Johnson. Why watch bees? motivations of citizen science volunteers in the great pollinator project. *Biological Conservation*, 208:40–47, 2017.

R.M. Edsall, L. Barbour, and J. Hoffman. Complementary methods for citizen mapping of ecosystem services: Comparing digital and analog representations. *Lecture Notes in Geoinformation and Cartography*, pages 295–307, 2015. ISSN 1863-2246. 27th International Cartographic Conference (ICC), Rio de Janeiro, BRAZIL, AUG 23-28, 2015.

Ria Follett et al. Exploring the actions of citizens as scientists through experimentation with aquaponics. 2015.

JOHN Francis, K Easterday, K Scheckel, and STEVEN R Beissinger. The world is a park: Using citizen science to engage people in parks and build the next century of global stewards. *Science, conservation, and national parks*, pages 275–293, 2017.

Amy Freitag. Factors determining participation quality in collaborative water quality research. *Environmental Sociology*, 3(3):248–259, 2017.

Didone Frigerio, Pavel Pipek, Sophia Kimmig, Silvia Winter, Jörg Melzheimer, Lucie Diblíková, Bettina Wachter, and Anett Richter. Citizen science and wildlife biology: Synergies and challenges. *Ethology*, 124(6):365–377, June 2018. ISSN 0179-1613.

Travis Gallo and Damon Waitt. Creating a successful citizen science model to detect and report invasive species. *BioScience*, 61(6):459–465, June 2011. ISSN 00063568, 15253244.

Mary M Gardiner, Leslie L Allee, Peter MJ Brown, John E Losey, Helen E Roy, and Rebecca Rice Smyth. Lessons from lady beetles: accuracy of monitoring data from us and uk citizen-science programs. *Frontiers in Ecology and the Environment*, 10(9):471–476, 2012.

Kelly Garvy. The emergence and use of angler self-reporting apps in recreational fisheries. 2015.

Carine Gimbert and Francois-Joseph Lapointe. Self-tracking the microbiome: where do we go from here? *MICROBIOME*, 3, dec 12 2015. ISSN 2049-2618.

Stefano Goffredo, Francesco Pensa, Patrizia Neri, Antonio Orlandi, Maria Scola Gagliardi, Angela Velardi, Corrado Piccinetti, and Francesco Zaccanti. Unite research with what citizens do for fun:'recreational monitoring' of marine biodiversity. *Ecological Applications*, 20 (8):2170–2187, 2010.

Washington Sea Grant, Don Meehan, Susan Bullerdick, COSEE-Ocean Learning Communities, Walter Pacheco, Muckleshoot Indian Tribe, Mike Racine, and Washington Scuba Alliance. Harnessing citizen science to protect and restore puget sound. 2009.

S. Gray and S. Scyphers. *Innovations in Collaborative Science: Advancing Citizen Science, Crowdsourcing and Participatory Modeling to Understand and Manage Marine Social-Ecological Systems*. 2017.

Bin Guan. The roles of organizations in environmental citizen science-cases from environmental monitoring in the uk. 2016.

S.A. Hamer, R. Curtis-Robles, and G.L. Hamer. Contributions of citizen scientists to arthropod vector data in. *Current Opinion in Insect Science*, 28:98–104, 2018.

Courtney Hann. Citizen science research: A focus on historical whaling data and a current marine mammal citizen science project, whale mapp. 2015.

Courtney H Hann, Lei Lani Stelle, Andrew Szabo, and Leigh G Torres. Obstacles and opportunities of using a mobile app for marine mammal research. *ISPRS International Journal of Geo-Information*, 7(5):169, 2018.

Annamarie Hatcher. Citizen science in the bras d'or lake biosphere–hatcher.

P.B. Heidorn. Biodiversity and biocomplexity informatics: Policy and implementation science versus citizen science. pages 362–364, 2002.

B. Hollow, P.E.J. Roetman, M. Walter, and C.B. Daniels. Citizen science for policy development: The case of koala management in south australia. *Environmental Science and Policy*, 47:126–136, March 2015. ISSN 1462-9011.

J. Hunter and C.-H. Hsu. Formal acknowledgement of citizen scientists' contributions via dynamic data citations. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 9469:64–75, 2015.

Kieran Hyder, Bryony Townhill, Lucy G Anderson, Jane Delany, and John K Pinnegar. Can citizen science contribute to the evidence-base that underpins marine policy? *Marine policy*, 59:112–120, 2015.

J.R. Jambeck and K. Johnsen. Citizen-based litter and marine debris data collection and mapping. *Computing in Science and Engineering*, 17(4):20–26, 2015.

Nicholas E Johnson and François Grey. Landfill hunter: Learning about waste through public participation. *Human Computation*, 3(1):243–252, 2016.

Saket Singh Kaurav et al. Citizen science: A roadmap for science centres and museums. *Journal of Scientific Temper (JST)*, 3(1 & 2), 2015.

Steve Kelling, Alison Johnston, Wesley M Hochachka, Marshall Iliff, Daniel Fink, Jeff Gerbracht, Carl Lagoze, Frank A La Sorte, Travis Moore, Andrea Wiggins, et al. Can observation skills of citizen scientists be estimated using species accumulation curves? *PloS one*, 10(10):e0139600, 2015.

Rachel Kelly, Aysha Fleming, Gretta Pecl, Anett Richter, and Aletta Bonn. Social licence through citizen science: A tool for marine conservation. *bioRxiv*, page 266692, 2018.

A.R. Khan and M. Mishra. The tree count: An approach of green indexing in urban areas using citizen science for earth observation and geographic information system (gis). volume 12, pages 10075–10079, 2012.

Hiromi Kobori, Janis L Dickinson, Izumi Washitani, Ryo Sakurai, Tatsuya Amano, Naoya Komatsu, Wataru Kitamura, Shinichi Takagawa, Kazuo Koyama, Takao Ogawara, et al. Citizen science: a new approach to advance ecology, education, and conservation. *Ecological Research*, 31(1):1–19, January 2016. ISSN 0912-3814.

Julian Koch and Simon Stisen. Citizen science: A new perspective to advance spatial pattern evaluation in hydrology. *PloS one*, 12(5):e0178165, may 30 2017. ISSN 1932-6203.

G. Kragh, R. Stafford, S. Curtin, and A. Diaz. Environmental volunteer well-being: Managers' perception and actual well-being of volunteers. *F1000Research*, 5, 2016.

Caroline Kyi, Nicole Tse, and Sandra Khazam. The potential role of citizen conservation in re-shaping approaches to murals in an urban context. *Studies in Conservation*, 61(sup2):98–103, 2016. ISSN 0039-3630.

Jeffrey Laut, Francesco Cappa, Oded Nov, and Maurizio Porfiri. Increasing patient engagement in rehabilitation through citizen science. In *ASME 2014 Dynamic Systems and Control Conference*, volume 2, pages V002T16A003–V002T16A003. American Society of Mechanical Engineers, 2014a. ISBN 978-0-7918-4619-3. ASME 7th Annual Dynamic Systems and Control Conference, San Antonio, TX, OCT 22-24, 2014.

Jeffrey Laut, Emiliano Henry, Oded Nov, and Maurizio Porfiri. Development of a mechatronics-based citizen science platform for aquatic environmental monitoring. *IEEE/ASME Transactions on Mechatronics*, 19(5):1541–1551, October 2014b. ISSN 1083-4435.

B. Lawson, S.O. Petrovan, and A.A. Cunningham. Citizen science and wildlife disease surveillance. *EcoHealth*, 12(4):693–702, December 2015. ISSN 1612-9202.

Adam S Lerner and Pat J Gehrke. Engaging public ecologies. In *Organic Public Engagement*, pages 39–67. Springer, 2018.

Roman Lukyanenko and Jeffrey Parsons. Using field experimentation to understand information quality in user-generated content. *CodeCon MIT*, 2014.

ROMAN LUKYANENKO, JEFFREY PARSONS, and YOLANDA F WIERSMA. Finding value through instance-based data collection in citizen science.

Soledad Luna, Margaret Gold, Alexandra Albert, Luigi Ceccaroni, Bernat Claramunt, Olha Danylo, Muki Haklay, Renzo Kottmann, Christopher Kyba, Jaume Piera, et al. Developing mobile applications for environmental and biodiversity citizen science: considerations and recommendations. In *Multimedia Tools and Applications for Environmental & Biodiversity Informatics*, pages 9–30. Springer, 2018.

Renée Lyons, Cassie F Quigley, and Michelle Cook. Care-based citizen science: Nurturing an ethic of care to support the preservation of biodiversity. In *Animals and Science Education*, pages 201–222. Springer, 2017.

Bruce J. MacFadden, Lisa Lundgren, Kent Crippen, Betty A. Dunckel, and Shari Ellis. Amateur paleontological societies and fossil clubs, interactions with professional paleontologists, and social paleontology in the united states. *PALAEONTOLOGIA ELECTRONICA*, 19(2), 2016. ISSN 1935-3952.

Susan F Magdziarz. *Examining participation in a Dolphin Observation Citizen Science program*. California State University, Long Beach, 2013.

P.G. Mahaffy, B. Martin, K.J. Ooms, A.F. Tappenden, M. Oliver, R. Hislop-Hook, J.E. Forman, U. Mans, and J. Sabou. Citizen science and international collaboration through environmental monitoring with simple chemical sensors. *Pure and Applied Chemistry*, 89(2):221–229, February 2017. ISSN 0033-4545.

Thomas Maillart, Mingyi Zhao, Jens Grossklags, and John Chuang. Given enough eyeballs, all bugs are shallow? revisiting eric raymond with bug bounty programs. *Journal of Cybersecurity*, 3(2): 81–90, 2017.

Elizabete Marchante and Helia Marchante. Engaging society to fight invasive alien plants in portugal-one of the main threats to biodiversity. In P Castro, UM Azeiteiro, P BacelarNicolau, WL Filho, and AM Azul, editors, *BIODIVERSITY AND EDUCATION FOR SUSTAINABLE DEVELOPMENT*, World Sustainability Series, pages 107–122. 2016. ISBN 978-3-319-32318-3; 978-3-319-32317-6.

H. Marchante, M.C. Morais, A. Gamela, and E. Marchante. Using a webmapping platform to engage volunteers to collect data on invasive plants distribution. *Transactions in GIS*, 21(2):238–252, April 2017. ISSN 1361-1682.

Victoria Martin, Liam Smith, Alison Bowling, Les Christidis, David Lloyd, and Gretta Pecl. Citizens as scientists: what influences public contributions to marine research? *Science Communication*, 38(4):495–522, August 2016a. ISSN 1075-5470.

V.Y. Martin, L. Christidis, and G.T. Pecl. Public interest in marine citizen science: Is there potential for growth? *BioScience*, 66(8):683–692, August 2016b. ISSN 0006-3568.

Alison M. Meadow, Michael A. Crimmins, and Daniel B. Ferguson. Field of dreams or dream team? assessing two models for drought impact reporting in the semiarid southwest. *BULLETIN OF THE AMERICAN METEOROLOGICAL SOCIETY*, 94(10):1507–1517, October 2013. ISSN 0003-0007.

P.A. Mieras, C. Harvey-Clark, M. Bear, G. Hodgin, and B. Hodgin. The economy of shark conservation in the northeast pacific: The role of ecotourism and citizen science. *Advances in Marine Biology*, 78:121–153, 2017. ISSN 0065-2881.

Edward E Millar, EC Hazell, and SJ Melles. The 'cottage effect'in citizen science? spatial bias in aquatic monitoring programs. *International Journal of Geographical Information Science*, pages 1–21, 2018.

R. Milne and B. Hansen. Engaging communities for prioritising natural resource management and biodiversity conservation actions. volume 1570, pages 36–40, 2016.

Y. Miyazaki, A. Murase, R. Sahara, A. Angulo, and H. Senou. Adding fish images taken in other countries to the biodiversity database of a japanese public museum, with report of range extension of labrisomus jenkinsi from the pacific coast of costa rica. *Ecological Research*, 32(1):89–93, January 2017. ISSN 0912-3814.

Michael J O'Grady, Conor Muldoon, Dominic Carr, Jie Wan, Barnard Kroon, and Gregory MP O'Hare. Intelligent sensing for citizen science. *Mobile Networks and Applications*, 21(2):375–385, 2016.

Monica A. Peters, David Hamilton, Chris Eames, John Innes, and Norman W.H. Mason. The current state of community-based environmental monitoring in new zealand. *New Zealand Journal of Ecology*, 40(3):279–288, 2016. ISSN 01106465, 11777788.

Christopher Phillips, Dylan Walshe, Karen O'Regan, Ken Strong, Christopher Hennon, Ken Knapp, Conor Murphy, and Peter Thorne. Assessing citizen science participation skill for altruism or university course credit: a case study analysis. 2018.

R. Phillips, Y. Ford, K. Sadler, S. Silve, and S. Baurley. Open design: Non-professional user-designers creating products for citizen science: A case study of beekeepers. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 8015 LNCS(PART 4):424–431, 2013.

Robert Phillips and Sharon Baurley. Exploring open design for the application of citizen science; a toolkit methodology. *Design Research Society*, 10(xx):1–15, 2014.

Robert Phillips, Michael Brown, and Sharon Baurley. Social responses to nature; citizen empowerment through design. *Journal of Design, Business & Society*, 2(2):197–215, 2016.

Michael JO Pocock, Helen E Roy, Richard Fox, Willem N Ellis, and Marc Botham. Citizen science and invasive alien species: predicting the detection of the oak processionary moth thaumetopoea processionea by moth recorders. *Biological conservation*, 208:146–154, 2017a.

Michael JO Pocock, Helen E Roy, Chris D Preston, and David B Roy. The biological records centre: a pioneer of citizen science. *Biological Journal of the Linnean Society*, 115(3):475–493, July 2015. ISSN 0024-4066.

Michael JO Pocock, John C Tweddle, Joanna Savage, Lucy D Robinson, and Helen E Roy. The diversity and evolution of ecological and environmental citizen science. *PloS one*, 12(4):e0172579, 2017b.

Jenny Preece, Kazjon Grace, Carol Boston, Mary Lou Maher, Tom Yeh, and Abigale Stangl. Crowdsourcing design and citizen science data using a tabletop in a nature preserve. *ECSM 2014 University of Brighton Brighton, UK 10-11 July 2014*, page 413, 2014.

D.G. Rossiter, J. Liu, S. Carlisle, and A.-X. Zhu. Can citizen science assist digital soil mapping? *Geoderma*, 259-260: 71–80, December 2015. ISSN 0016-7061.

Helen E Roy, Michael JO Pocock, Chris D Preston, David B Roy, J Savage, JC Tweddle, and LD Robinson. Understanding citizen science and environmental monitoring: final report on behalf of uk environmental observation framework. 2012.

Shana Alysse Sandhaus. Evaluating the motivations, knowledge, and efficacy of participants in environmental health citizen science projects. 2017.

L.D. Savio, B. Prainsack, and A. Buyx. Motivations of participants in the citizen science of microbiomics: Data from the british gut project. *Genetics in Medicine*, 19(8):959–961, 2017.

S.B. Scyphers, S.P. Powers, J.L. Akins, J.M. Drymon, C.W. Martin, Z.H. Schobernd, P.J. Schofield, R.L. Shipp, and T.S. Switzer. The role of citizens in detecting and responding to a rapid marine invasion. *Conservation Letters*, 8(4):242–250, jul-aug 2015. ISSN 1755-263X.

VI Seymour and Mordechai Haklay. Exploring engagement characteristics and behaviours of environmental volunteers. *Citizen Science: Theory and Practice*, 2 (1), 2017.

S Sheppard. Managing data quality in observational citizen science. 2017.

Sergei Yu Shevchenko. Citizen science: are people distinguishable from bacteria? *Epistemology & Philosophy of Science*, 55(1):171–183, 2018.

Ulrike Sturm, Alexandra Moormann, and Astrid Faber. Mobile learning in environmental citizen science: An initial survey of current practice in germany. *it-Information Technology*, 60(1):3–9, 2018.

William J Sutherland, David B Roy, and Tatsuya Amano. An agenda for the future of biological recording for ecological monitoring and citizen science. *Biological Journal of the Linnean Society*, 115(3):779–784, 2015.

Sonja Teichert. The influence of capacity and attitudes in the use of water quality citizen science and volunteer benthic monitoring in the freshwater management activities of ontario's conservation authorities. 2016.

Jan Top and Mari Wigham. 5 the role of e-science in agriculture: How e-science technology assists participation in agricultural research. *KNOWLEDGE AND INNOVATION SYSTEMS TOWARDS THE FUTURE*, page 45.

Ginger Tsueng, Max Nanis, Jennifer Fouquier, Benjamin Good, and Andrew Su. Citizen science for mining the biomedical literature. *bioRxiv*, page 38083, 2016.

Markel Vigo, Lamiece Hassan, William Vance, Caroline Jay, Andrew Brass, and Sheena Cruickshank. Britain breathing: using the experience sampling method to collect the seasonal allergy symptoms of a country. *Journal of the American Medical Informatics Association*, 25(1): 88–92, 2017.

E.D. Villasclaras-Fernandez, M. Sharples, S. Kelley, and E. Scanlon. Supporting citizen inquiry: An investigation of moon rock. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 8095 LNCS:383–395, 2013. ISSN 0302-9743. 8th European Conference on Technology Enhanced Learning (EC-TEL), Paphos, CYPRUS, SEP 17-21, 2013.

V. Visser, B. Langdon, A. Pauchard, and D.M. Richardson. Unlocking the potential of google earth as a tool in invasion science. *Biological Invasions*, 16(3):513–534, March 2014. ISSN 1387-3547.

H Volten, JLA Devilee, A Apituley, LJ Carton, M Grothe, C Keller, F Kresin, A Land-Zandstra, E Noordijk, E van Putten, et al. Citizen science with small sensor networks. collaboration between a dutch epa (rivm) and local initiatives. 2016.

Susan WEST, Darlene CAVALIER, and Michael GOLD. Citizen science from the citizen's point of view. In *2010 GSA Denver Annual Meeting*, 2010.

A. Wiggins. ebirding: Technology adoption and the transformation of leisure into science. pages 798–799, 2011.

R.L. Williams, R. Stafford, and A.E. Goodenough. Biodiversity in urban gardens: Assessing the accuracy of citizen science data on garden hedgehogs. *Urban Ecosystems*, 18(3):819–833, September 2015. ISSN 1083-8155.

Debbie Winton, Mark Huxham, and Jenny A Cousins. Using citizen science to address conservation issues related to climate change and coastal systems. In *Citizen Science for Coastal and Marine Conservation*, pages 39–58. Routledge, 2017.

HaoFan Yang, Jinglan Zhang, and Paul Roe. Reputation modelling in citizen science for environmental acoustic data analysis. *Social Network Analysis and Mining*, 3(3):419–435, 2013.

J. Yu, W.-K. Wong, and R.A. Hutchinson. Modeling experts and novices in citizen science data for species distribution modeling. pages 1157–1162, 2010.

A.8 Other Languages

Articles which are not available in English:

A. Alonso-Puelles. What we learned from aids and electro-sensitivity: Vulnerable communities and their empowerment [quÃ aprendimos del sida y de la electro-sensibilidad: Comunidades vulnerables y su empoderamiento]. *Profesional de la Informacion*, 26(1): 106–113, jan-feb 2017. ISSN 1386-6710.

Hans-Günther Bauer, Olaf Geiter, Susanne Homma, and Friederike Woog. Die neuen statuseinstufungen der fachgruppe neozoen der do-g: Entgegnung zu fritz et al.(2017): Aspekte der nahrungsökologie und genetik des waldrapps. *Redaktion/Schriftleitung*, page 146, 2017.

André Luiz Brazil and Sarita Albagli. Usos da" gamificação" na produção colaborativa de informação e conhecimento. 2017.

Suzanne Buchan, Andres Janser, et al. *Animierte Wunderwelten/Animated Wonderworlds*. Schwabe AG, 2015.

Eva Bunge. *Citizen Science in der Bibliotheksarbeit: Möglichkeiten und Chancen*, volume 63. bit online Verlag, 2017.

Emanuela Chiriac. Wikipédia, la chimère du savoir libre. *Documentation et bibliothèques*, 61(4):159–166, 2015.

Lisa Chupin. L'informatisation participative d'herbiers comme dispositif de médiation du patrimoine et de la culture scientifique. quelles convergences avec les dispositifs de formation à la médiation? *Tréma*, (48):119–131, 2018.

Jorge García del Arco. Plataforma de crowdsourcing con un modelo conceptual semántico para la optimización en el proceso de asignación de.

P. Dias Da Silva, L. Heaton, and F. Millerand. A review of the citizen science literature: Producing naturalist knowledge in the digital age [une revue de littÃrature sur la Â science citoyenne Â≫: la production de connaissances naturalistes Ã l'Ã¨re numÃrique?]. *Natures Sciences Societes*, 25(4):370–380, 2017.

S. Dickel and M. Franzen. Digital inclusion: The social implications of open science [digitale inklusion: Zur sozialen Ãffnung des wissenschaftssystems]. *Zeitschrift fur Soziologie*, 44(5):330–347, October 2015. ISSN 0340-1804.

A. Dippel. The big data game: On the ludic constitution of the collaborative production of knowledge in high-energy physics at cern [das big data game: Zur spielerischen konstitution kollaborativer wissensproduktion in der hochenergiephysik am cern]. *NTM International Journal of History and Ethics of Natural Sciences, Technology and Medicine*, 25 (4):485–517, December 2017a. ISSN 0036-6978.

Anne Dippel. Das big data game. *NTM Zeitschrift für Geschichte der Wissenschaften, Technik und Medizin*, 25(4): 485–517, 2017b.

CMIA de Vila do Conde. Ea, interpretación e conservación.

Iker Etxebarria, Alicia García, and Jorge García del Arco. Plataforma de crowdsourcing con un modelo conceptual semántico para la optimización en el proceso de asignación de recursos microfinancieros. *CONSEJO EDITORIAL*, page 39.

M. Fejes, M.G.P. Sales, and M.E. Infante-Malachias. Using simulations as a different dynamic to learners and teachers: A case study. *Journal of Science Education*, 14(SPEC. ISSUE):7–11, 2013.

Ivan Roberto Ferraz, Maria Aparecida Gouvêa, and Iná Futino Barreto. Fatores determinantes da participação em iniciativas de crowdsourcing. *Revista Gestão & Tecnologia*, 17(1):107–129, 2017.

I. Franz, C.E. Agne, G.A. Bencke,

L. Bugoni, and R.A. Dias. Four decades after belton: A review of records and evidences on the avifauna of rio grande do sul, brazil [quatro dÃcadas apÃs belton: Uma revisÃ£o de registros e evidÃªncias sobre a avifauna do rio grande do sul, brasil]. *Iheringia - Serie Zoologia*, 108, 2018. ISSN 0073-4721.

Von BIRGIT GANTNER, BRIGITTE ALLE, CHRISTIANE BRANDENBURG, JULIA KELEMEN-FINAN, THOMAS HOLZER, SILVIA WINTER, MONIkA KRIECHBAUM, MARGIT SEIBERL, and URSULA LIEBL. Laienmonitoring mit schülern.

Okke Gerhard, Nils Wolf, and Alexander Siegmund. Einsatz von citizen science im phänologischen monitoring der apfelblüte in deutschland. *Wissenschaft für alle: Citizen Science*, page 123, 2017.

Inge Heyer, Stephanie J Slater, and Timothy F Slater. Estabelecendo uma relação empírica entre o raciocínio espacial dos estudantes de graduação em carreiras não científicas e seu conhecimento conceitual da astronomia. *Revista Latino-Americana de Educação em Astronomia*, (16):45–61, 2013.

Ingrid Hjorth-Johansen. Piggsvin (erinaceus europaeus) i oslo: vurdering av metoder for kartlegging, inkludert bruk av sosiale medier, samt betydning av arealbruk for arten. Master's thesis, Norwegian University of Life Sciences, Ås, 2016.

F. Houllier, P.-B. Joly, and J.-B. Merilhou-Goudard. Citizen sciences: A dynamics to be encouraged [les sciences participatives: une dynamique Ã conforter?]. *Natures Sciences Societes*, 25(4):418–423, 2017.

Jessica Leander and Jakob Traung. 'det är inte rocket science att göra science'. en kvalitativ studie om hur citizen science-plattformar organiserar sig för att motivera till deltagande. 2015.

M. Legrand. Vigie-nature: Citizen science and biodiversity on a large scale [vigie-nature: Sciences participatives et biodiversitÃ Ã grande Ãchelle]. *Cahiers des Ameriques Latines*, (72-73):65–83, 2013.

G. Neuwirth and N. Hirneisen. Operating a web portal for field observations - experiences and challenges of the example 'www.naturbeobachtung.de' [betreiben eines onlineportals fÃr naturbeobachtungen- erfahrungen und herausforderungen am beispiel von www.naturbeobachtung.at]. *Naturschutz und Landschaftsplanung*, 45(6):177–182, 2013.

A. Orth and B. Schmidt. Open science lernen und lehren: Foster portal stellt materialien und kurse bereit open science education: Learning and teaching materials via the foster portal apprendre et enseigner l'open science : le portail foster fournit du matÃriel et des cours. *Information-Wissenschaft und Praxis*, 66(2-3):121–128, April 2015. ISSN 1434-4653.

Adriana Landim Quinaud and Maria José Baldessar. A educação no século xxi: gamificação aprendizagem com criatividade. *Temática*, 13(11), 2017.

Shirley Luana Ramos, Luıs Fabrıcio W Góes, Lesandro Ponciano, Celso França, and Hugo Morais. Avaliac ao da percepc ao de jogadores sobre a criatividade de combos do jogo digital de cartas hearthstone.

Anett Richter, Tabea Turrini, Karin Ulbrich, Anika Mahla, and Aletta Bonn. Citizen science–perspektiven in der umweltbildung.

IE Salverda, PD van der Jagt, R Willemse, MC Onwezen, and JL Top. Sociale media: nieuwe wegen naar sociale innovatie. Technical report, 2013.

Daniel Schneider, Laure Kloetzer, and Julien DaCosta. Apprendre en participant à des projets ≪citizen science≫ numériques. *Raisons éducatives*, (1):229–248, 2017.

Jeroen Speybroeck, Anton Stumpel, Wouter Beukema, Bobby Bok, Raymond Creemers, Jeroen van Delft, Henk Strijbosch, and Jan Van Der Voort. Standaardlijst van nederlandse namen van de europese amfibieën en reptielen-uptdate naar situatie 2016. *RAVON*, 18 (4):71–76, 2016.

Vidal Torres. R.(2015). les compétences des participants dans les sciences participatives sur internet: une exploration de foldit. *TEM*, pages 3–5, 2015.

C Van Swaay, E Regan, M Ling, E Bozhinovska, M Fernandez, OJ Marini-Filho, B Huertas, and CK Phon. K' orösi. *A., Meerman, J., Pe'er, G., Uehara-Prado, M., Sáfián, S., Sam, L., Shuey, J., Taron, D., Terblanche, R., and Underhill, L*, 2015.

C. Von Aden, F. Kastner, J. Loesbrock, and S. Krohn-Grimberghe. New approaches for detecting digital modes the voluntary nature conservation results of the development of mobile solutions in lower saxony [neue ansÃtze digitaler artenerfassung fÃr den ehrenamtlichen naturschutz:ergebnisse der entwicklung mobiler lÃ¶sungen in niedersachsen]. *Naturschutz und Landschaftsplanung*, 45 (4):101–107, 2013.

Gert G Wagner, Michaela Engelmann, Jan Goebel, Florian Griese, Marcel Hebing, Janine Napieraj, Marius Pahl, Carolin Stolpe, Monika Wimmer, Alexander Eickelpasch, et al. Citizen science" auf basis des soep: Entwicklung und erste anwendung eines software-tools für" bürgerdialoge. Technical report, SOEPpapers on Multidisciplinary Panel Data Research, 2014.

Y.F. Wiersma. Birding 2.0: Citizen science and effective monitoring in the web 2.0 world [ornithologie 2.0: La science citoyenne et les programmes de suivi Ã l'Ã¨re d'internet 2.0]. *Avian Conservation and Ecology*, 5(2), December 2010. ISSN 1712-6568.

T. Wohlgemuth, A. Nussbaumer, A. Burkart, M. Moritzi, U. Wasem, and B. Moser. Patterns and driving forces for seed production in forest tree species [muster und treibende krÃfte der samenproduktion bei waldbÃumen]. *Schweizerische Zeitschrift fur Forstwesen*, 167(6):316–324, 2016.

M. Zacklad and L. Chupin. Scientific and heritage crowdsourcing at the crossroads of models of coordination and cooperation: The case of digital herbaria [le crowdsourcing scientifique et patrimonial Ã la croisÃe de modÃ¨les de coordination et de coopÃration: Le cas des herbiers numÃrisÃs]. *Canadian Journal of Information and Library Science*, 39(3-4):308–328, sep-dec 2015. ISSN 1195-096X.

A.9 Other Related Concepts

Articles which describe other concepts which are similar to (but not) VCS, such as Volunteer Geographic Information:

Christian Bröer, Gerben Moerman, Johan Casper Wester, Liza Rubinstein Malamud, Lianne Schmidt, Annemiek Stoopendaal, Nynke Kruiderink, Christina Hansen, and Hege Sjølie. Open online research: Developing software and method for collaborative interpretation. 2016.

C Christian, G Gerben, Johan Casper JC, Liza Rubinstein LR, L Lianne, Annemiek Stoopendaal, N Nynke, C Christina, and H Hege. Open online research: Developing software and method for collaborative interpretation. In *Forum Qualitative Sozialforschung*, volume 17, 2016.

Evgenia Christoforou. Achieving reliability and fairness in online task computing environments. 2017.

A. Delfanti. Users and peers. from citizen science to p2p science. *Journal of Science Communication*, 9(1):1–5, 2010.

S. Dickel and M. Franzen. The "problem of extension" revisited: New modes of digital participation in science. *Journal of Science Communication*, 15(1), 2016. ISSN 1824-2049.

Sascha Dickel. Trust in technologies? science after de-professionalization. *J Sci Commun*, 15:1–7, 2016.

Heather Ford. The person in the (big) data: A selection of innovative methods, strategies and perspectives for social research in the age of (big) data. *Working Papers of the Communities & Culture Network+*, 8, 2016.

Michael Nielsen. *Reinventing discovery: the new era of networked science.* Princeton University Press, 2011.

Roberto CS Pacheco. Digital science: Cyberinfrastructure, e-science and citizen science roberto cs pacheco; everton nascimento and rosina o. weber. pages 377–388, 2018.

B. Pirenne and E. Guillemot. Beyond data management: How to foster data exploitation and better science? In *2012 OCEANS*. IEEE; Marine Technol Soc (MTS); IEEE OES (IEEE/OES), 2012. ISBN 978-1-4673-0829-8. MTS/IEEE Oceans Conference, Virginia Beach, VA, OCT 14-19, 2012.

Eileen Scanlon, Mike Sharples, and Christothea Herodotou. Introducing citizen inquiry. In *Citizen Inquiry*, pages 19–24. Routledge, 2017.

A.10 Results

Publications which describe results from Citizen Science projects:

Jeffrey Ackley. *Rich Lizards: How Affluence, Land Cover, and the Urban Heat Island Effect Influence Desert Reptiles Persisting in an Urban Landscape*. Arizona State University, 2015.

F. BÃthori, W.P. Pfliegler, C.-U. Zimmerman, and A. Tartally. Online image databases as multi-purpose resources: Discovery of a new host ant of rickia wasmannii cavara (ascomycota, laboulbeniales) by screening antweb.org. *Journal of Hymenoptera Research*, 61: 85–94, dec 20 2017. ISSN 1070-9428.

L. Barnard, C. Scott, M. Owens, M. Lockwood, K. Tucker-Hood, S. Thomas, S. Crothers, J.A. Davies, R. Harrison, C. Lintott, R. Simpson, J. O'Donnell, A.M. Smith, N. Waterson, S. Bamford, F. Romeo, M. Kukula, B. Owens, N. Savani, J. Wilkinson, E. Baeten, L. Poeffel, and B. Harder. The solar stormwatch cme catalogue: Results from the first space weather citizen science project. *Space Weather*, 12(12):657–674, 2014.

E.F. Beach, M. Gilliver, and W. Williams. Hearing protection devices: Use at work predicts use at play. *Archives of Environmental and Occupational Health*, 71(5):281–288, 2016. ISSN 1933-8244.

Stephanie Bird and Joel Parker. Low levels of light pollution may block the ability of male glow-worms (lampyris noctiluca l.) to locate females. *Journal of insect conservation*, 18(4):737–743, 2014.

Jeffrey M. Black. River otter monitoring by citizen science volunteers in northern california: Social groups and litter size. *Northwestern Naturalist*, 90(2):130–135, 2009. ISSN 10511733, 19385315.

Alexander L Bond and Jennifer L Lavers. Flesh-footed shearwaters (puffinus carneipes) in the northeastern pacific ocean: summary and synthesis of records from canada and alaska. *The Canadian Field-Naturalist*, 129(3):263–267, 2015.

Ronald J. Buta. Galactic rings revisited - i. cvrhs classifications of 3962 ringed galaxies from the galaxy zoo 2 database. *MONTHLY NOTICES OF THE ROYAL ASTRONOMICAL SOCIETY*, 471(4):4027–4046, November 2017. ISSN 0035-8711.

A. Campanaro, S. Hardersen, L.R.D. Zan, G. Antonini, M. Bardiani, M. Maura, E. Maurizi, F. Mosconi, A. Zauli, M.A. Bologna, P.F. Roversi, G.S. Peverieri, and F. Mason. Analyses of occurrence data of protected insect species collected by citizens in italy. *Nature Conservation*, 20(20, SI):265–297, 2017. ISSN 1314-6947.

M. Caruana. Analysis of data from maltese passport applications held at the national archives of malta: A new digital resource. *New Review of Information Networking*, 21(1):52–62, 2016.

Giorgos Chatzigeorgiou, Sarah Faulwetter, Thanos Dailianis, Vincent Stuart Smith, Panagiota Koulouri, Costas Dounas, and Christos Arvanitidis. Testing the robustness of citizen science projects: Evaluating the results of pilot project comber. *Biodiversity data journal*, (4), 2016.

Charlotte Duranton, Thierry Bedossa, and Florence Gaunet. The perception of dogs' behavioural synchronization with their owners depends partially on expertise in behaviour. *Applied Animal Behaviour Science*, 199:24–28, 2018.

G. Ehmke, G.S. Maguire, T. Bird, D. Ierodiaconou, and M.A. Weston. An obligate beach bird selects sub-, inter- and supra-tidal habitat elements. *Estuarine, Coastal and Shelf Science*, 181: 266–276, nov 5 2016. ISSN 0272-7714.

J. Everaars, M.W. Strohbach, B. Gruber, and C.F. Dormann. Microsite conditions dominate habitat selection of the red mason bee (osmia bicornis, hymenoptera: Megachilidae) in an urban environment: A case study from leipzig, germany. *Landscape and Urban Planning*, 103(1):15–23, oct 30 2011. ISSN 0169-2046.

Daniel Fink, Wesley M Hochachka, and Steve Kelling. Cos 103-8: Modeling species distribution dynamics with spatiotemporal exploratory models: Discovering patterns and processes of broad-scale avian migrations. In *The 95th ESA Annual Meeting*, 2010.

Daniel Fink, Wesley M Hochack, Benjamin Zuckerberg, and Steve T Kelling. Modeling species distribution dynamics with spatiotemporal exploratory models: Discovering patterns and processes of broad-scale avian migrations. *Procedia Environmental Sciences*, 7:50–55, 2011. ISSN 1878-0296. 1st International Conference on Spatial Statistics - Mapping Global Change, Enschede, NETHERLANDS, MAR 23-25, 2011.

R Fox, TM Brereton, J Asher, TA August, MS Botham, NAD Bourn, KL Cruickshanks, CR Bulman, S Ellis, CA Harrower, et al. The state of the uk's butterflies 2015. 2015.

J. Gago, P. Anastacio, C. Gkenas, F. Banha, and F. Ribeiro. Spatial distribution patterns of the non-native european catfish, silurus glanis, from multiple online sources - a case study for the river tagus (iberian peninsula). *FISHERIES MANAGEMENT AND ECOLOGY*, 23(6):503–509, December 2016. ISSN 0969-997X.

Melanie A Galloway, Kyle W Willett, Lucy F Fortson, Carolin N Cardamone, Kevin Schawinski, Edmond Cheung, Chris J Lintott, Karen L Masters, Thomas Melvin, and Brooke D Simmons. Galaxy zoo: the effect of bar-driven fuelling on the presence of an active galactic nucleus in disc galaxies. *Monthly Notices of the Royal Astronomical Society*, 448(4): 3442–3454, 2015.

J. E. Geach, Y. T. Lin, M. Makler, J. P. Kneib, N. P. Ross, W. H. Wang, B. C. Hsieh, A. Leauthaud, K. Bundy, H. J. McCracken, J. Comparat, G. B. Caminha, P. Hudelot, L. Lin, L. Van Waerbeke, M. E. S. Pereira, and D. Mast. Vics82: The vista-cfht stripe 82 near-infrared survey. *ASTROPHYSICAL JOURNAL SUPPLEMENT SERIES*, 231 (1), July 2017. ISSN 0067-0049.

M.F. Geiger, J.J. Astrin, T. Borsch, U. Burkhardt, P. Grobe, R. Hand, A. Hausmann, K. Hohberg, L. Krogmann, M. Lutz, C. Monje, B. Misof, J. MoriniÃ¨re, K. MÃller, S. Pietsch, D. Quandt, B. Rulik, M. Scholler, W. Traunspurger, G. Haszprunar, and W. WÃgele. How to tackle the molecular species inventory for an industrialized nation-lessons from the first phase of the german barcode of life initiative gbol (2012-2015). *Genome*, 59(9):661–670, September 2016. ISSN 0831-2796.

Rose A Graves, Scott M Pearson, and Monica G Turner. –bird community dynamics change the seasonal distribution of a cultural ecosystem service in a montane landscape. *SPATIAL DYNAMICS OF BIODIVERSITY-BASED ECOSYSTEM SERVICES IN THE SOUTHERN APPALACHIAN MOUNTAINS*, 1001:163.

Rose A Graves, Scott M Pearson, and Monica G Turner. Effects of bird community dynamics on the seasonal distribution of cultural ecosystem services. *Ambio*, pages 1–13, 2018.

Stuart Grayson. Counting weddell seals project report. 2017.

CJ Hansen, G Portyankina, KM Aye, ME Schwamb, C Lintott, and A Smith. Mars' seasonal fans measured by citizen scientists. In *European Planetary Science Congress*, volume 8, 2013.

M.D. Hellwig. Automatic time-series quantification of bluff erosion using a single consumer grade camera as basis for erosion risk assessment and forecasts – a boston harbor islands case study. *Journal of Coastal Conservation*, 20(6):469–476, December 2016. ISSN 1400-0350.

Robbert Hoeboer. Multi-scale approaches to identify topographic features favourable as breeding sites by malaria vectors. 2016.

Elizabeth Howard, Harlen Aschen,

and Andrew K Davis. Citizen science observations of monarch butterfly overwintering in the southern united states. *Psyche: A Journal of Entomology*, 2010, 2010.

Elizabeth Howard and Andrew K Davis. The fall migration flyways of monarch butterflies in eastern north america revealed by citizen scientists. *Journal of Insect Conservation*, 13(3):279–286, 2009.

Elizabeth Howard and Andrew K. Davis. *Tracking the Fall Migration of Eastern Monarchs with Journey North Roost Sightings: New Findings about the Pace of Fall Migration*, pages 207–214. Cornell University Press, 1 edition, 2015. ISBN 9780801453151.

Henry H. Hsieh, Larry Denneau, Richard J. Wainscoat, Norbert Schoerghofer, Bryce Bolin, Alan Fitzsimmons, Robert Jedicke, Jan Kleyna, Marco Micheli, Peter Veres, Nicholas Kaiser, Kenneth C. Chambers, William S. Burgett, Heather Flewelling, Klaus W. Hodapp, Eugene A. Magnier, Jeffrey S. Morgan, Paul A. Price, John L. Tonry, and Christopher Waters. The main-belt comets: The pan-starrs1 perspective. *ICARUS*, 248:289–312, mar 1 2015. ISSN 0019-1035.

Allen H Hurlbert and Zhongfei Liang. Spatiotemporal variation in avian migration phenology: citizen science reveals effects of climate change. *PLoS One*, 7 (2):e31662, 2012.

Tenley Hutchinson-Smith and Brooke Simmons. Bar evolution and bar properties from disc galaxies in the early universe. In *American Astronomical Society Meeting Abstracts*, volume 229, 2017.

Christopher Izzo, Zoë A Doubleday, Gretchen L Grammer, Kayla L Gilmore, Heidi K Alleway, Thomas C Barnes, Morgan CF Disspain, Ana Judith Giraldo, Nastaran Mazloumi, and Bronwyn M Gillanders. Fish as proxies of ecological and environmental change. *Reviews in fish biology and fisheries*, 26 (3):265–286, 2016.

Michelle M Jackson, Sarah E Gergel, and Kathy Martin. Citizen science and field survey observations provide comparable results for mapping vancouver island white-tailed ptarmigan (lagopus leucura saxatilis) distributions. *Biological Conservation*, 181:162–172, 2015.

Andrew JB Jennings and Stanley H Faeth. Preformed scour holes associated with road building may maintain anuran diversity in urbanizing areas.

L. Clifton Johnson, Anil C. Seth, Julianne J. Dalcanton, Matthew L. Wallace, Robert J. Simpson, Chris J. Lintott, Amit Kapadia, Evan D. Skillman, Nelson Caldwell, Morgan Fouesneau, Daniel R. Weisz, Benjamin F. Williams, Lori C. Beerman, Dimitrios A. Gouliermis, and Ata Sarajedini. Phat stellar cluster survey. ii. andromeda project cluster catalog. *ASTROPHYSICAL JOURNAL*, 802(2), apr 1 2015. ISSN 0004-637X.

Heather R Jordan and Jeffery K Tomberlin. Abiotic and biotic factors regulating inter-kingdom engagement between insects and microbe activity on vertebrate remains. *Insects*, 8(2):54, 2017.

S. Jouveau, M. Delaunay, R. Vignes-Lebbe, and R. Nattier. A multi-access identification key based on colour patterns in ladybirds (coleoptera, coccinellidae). *ZooKeys*, 2018(758):55–73, may 14 2018. ISSN 1313-2989.

Alison Ke and Matthew Scott Luskin. Integrating disparate occurrence reports to map data-poor species ranges and occupancy: a case study of the vulnerable bearded pig sus barbatus. *Oryx*, pages 1–11, 2017.

I. Kelly, J.X. Leon, B.L. Gilby, A.D. Olds, and T.A. Schlacher. Marine turtles are not fussy nesters: A novel test of small-scale nest site selection using structure from motion beach terrain information. *PeerJ*, 2017(1), jan 4 2017. ISSN 2167-8359.

Jeffrey F Kelly and Sandra M Pletschet. Accuracy of swallow roost locations assigned using weather surveillance radar. *Remote Sensing in Ecology and Conservation*, 2017.

Madeleine M Kern, Jacquelyn C Guzy, Steven J Price, Stephanie D Hunt, Evan A Eskew, and Michael E Dorcas. Riparian-zone amphibians and reptiles within the broad river basin of south carolina. *Journal of North Carolina Academy of Science*, 128(3):81–87, 2012.

Charles R Kerton, Grace Wolf-Chase, Kim Arvidsson, Chris J Lintott, and Rob J Simpson. The milky way project: What are yellowballs? *The Astrophysical Journal*, 799(2):153, 2015.

F. Khatib, S. Cooper, M.D. Tyka, K. Xu, I. Makedon, Z. PopoviÄ‡, D. Baker, and F. Players. Algorithm discovery by protein folding game players. *Proceedings of the National Academy of Sciences of the United States of America*, 108(47):18949–18953, nov 22 2011. ISSN 0027-8424.

Kees Koffijberg, Erik van Winden, and Preben Clausen. The netherlands as a winter refuge for light-bellied brent geese branta bernicla hrota. *Wildfowl*, pages 40–56, 2013.

Ondrej Korabek, Lucie Jurickova, and Adam Petrusek. Resurrecting helix straminea, a forgotten escargot with trans-adriatic distribution: first insights into the genetic variation within the genus helix (gastropoda: Pulmonata). *ZOOLOGICAL JOURNAL OF THE LINNEAN SOCIETY*, 171(1):72–91, May 2014. ISSN 0024-4082.

S.J. Kruk, C.J. Lintott, S.P. Bamford, K.L. Masters, B.D. Simmons, B. HÃu AŸler, C.N. Cardamone, R.E. Hart, L. Kelvin, K. Schawinski, R.J. Smethurst, and M. Vika. Galaxy zoo: Secular evolution of barred galaxies from structural decomposition of multiband images. *Monthly Notices of the Royal Astronomical Society*, 473(4):4731–4753, February 2018. ISSN 0035-8711.

Ana Laranja, Silva Morim, and Rosana M Afonso. The primary role of exploratory trails in environmental education-cmia of vila do conde. *AmbientalMENTE Sustentable*, 1(23-24): 193–198, 2018.

Lincoln R Larson, April L Conway, Sonia M Hernandez, and John P Carroll. Human-wildlife conflict, conservation attitudes, and a potential role for citizen science in sierra leone, africa. *Conservation and Society*, 14(3):205–217, 2016.

J. Le Coz, A. Patalano, D. Collins, N.F. GuillÃn, C.M. GarciÃ, G.M. Smart, J. Bind, A. Chiaverini, R. Le Boursicaud, G. Dramais, and I. Braud. Lessons learnt from recent citizen science initiatives to document floods in france, argentina and new zealand. In M Lang, F Klijn, and P Samuels, editors, *3RD EUROPEAN CONFERENCE ON FLOOD RISK MANAGEMENT (FLOODRISK 2016)*, volume 7 of *E3S Web of Conferences*, 2016. 3rd European Conference on Flood Risk Management (FLOODrisk), Lyon, FRANCE, OCT 17-21, 2016.

J. Lee, W. Kladwang, M. Lee, D. Cantu, M. Azizyan, H. Kim, A. Limpaecher, S. Yoon, A. Treuille, and R. Das. Rna design rules from a massive open laboratory. *Proceedings of the National Academy of Sciences of the United States of America*, 111(6):2122–2127, feb 11 2014. ISSN 0027-8424.

J. Lisjak, S. Schade, and A. Kotsev. Closing data gaps with citizen science? findings from the danube region. *ISPRS International Journal of Geo-Information*, 6(9), September 2017. ISSN 2220-9964.

S.B. Liu, B.S. Poore, R.J. Snell, A. Goodman, N.G. Plant, H.F. Stockdon, K.L.M. Morgan, and M.D. Krohn. Usgs icoast - did the coast change? designing a crisis crowdsourcing app to validate coastal change models. pages 17–20, 2014.

Yong Liu, Pratch Piyawongwisal, Sahil Handa, Liang Yu, Yan Xu, and Arjmand Samuel. Going beyond citizen data collection with mapster: a mobile+ cloud real-time citizen science experiment. In *e-Science Workshops (eScienceW), 2011 IEEE Seventh International Conference on*, pages 1–6. IEEE, 2011.

M. McCampbell, M. Schut, I. Van den Bergh, B. van Schagen, B. Vanlauwe, G. Blomme, S. Gaidashova, E. Njukwe, and C. Leeuwis. Xanthomonas wilt of banana (bxw) in central africa: Opportunities, challenges, and pathways for citizen science and ict-based control and prevention strategies. *NJAS - Wageningen Journal of Life Sciences*, 2018.

Kent P McFarland, John D Lloyd, and Spencer P Hardy. Density and habitat relationships of the endemic white mountain fritillary (boloria chariclea montinus)(lepidoptera: Nymphalidae). *Insects*, 8(2):57, 2017.

E.S. Minor, C.W. Appelt, S. Grabiner, L. Ward, A. Moreno, and S. Pruett-Jones. Distribution of exotic monk parakeets across an urban landscape. *Urban Ecosystems*, 15(4):979–991, 2012.

Anupreeta More, Aprajita Verma, Philip J. Marshall, Surhud More, Elisabeth Baeten, Julianne Wilcox, Christine Macmillan, Claude Cornen, Amit Kapadia, Michael Parrish, Chris Snyder, Christopher P. Davis, Raphael Gavazzi, Chris J. Lintott, Robert Simpson, David Miller, Arfon M. Smith, Edward Paget, Prasenjit Saha, Rafael Kueng, and Thomas E. Collett. Space warps-ii. new gravitational lens candidates from the cfhtls discovered through citizen science. *MONTHLY NOTICES OF THE ROYAL ASTRONOMICAL SOCIETY*, 455(2):1191–1210, jan 11 2016. ISSN 0035-8711.

B.M. Norman, J.A. Holmberg, Z. Arzoumanian, S.D. Reynolds, R.P. Wilson, D. Rob, S.J. Pierce, A.C. Gleiss, R. De La Parra, B. Galvan, D. Ramirez-Macias, D. Robinson, S. Fox, R. Graham, D. Rowat, M. Potenski, M. Levine, J.A. McKinney, E. Hoffmayer, A.D.M. Dove, R. Hueter, A. Ponzo, G. Araujo, E. Aca, D. David, R. Rees, A. Duncan, C.A. Rohner, C.E.M. Prebble, A. Hearn, D. Acuna, M.L. Berumen, A. VÃzquez, J. Green, S.S. Bach, J.V. Schmidt, S.J. Beatty, and D.L. Morgan. Undersea constellations: The global biology of an endangered marine megavertebrate further informed through citizen science. *BioScience*, 67(12):1029–1043, December 2017. ISSN 0006-3568.

RYAN P O'Donnell and ANDREW M Durso. Harnessing the power of a global network of citizen herpetologists by improving citizen science databases. *Herpetological Review*, 45(1):151–157, 2014.

H Dieter Oschadleus. Birds adopting weaver nests for breeding in africa. *Ostrich*, 89(2):131–138, 2018.

L. Patterson, R. Kalle, and C. Downs. A citizen science survey: perceptions and attitudes of urban residents towards vervet monkeys. *Urban Ecosystems*, 20 (3):617–628, June 2017. ISSN 1083-8155.

Kevin A Pimbblet and Peter C Jensen. The role of stellar mass and environment for cluster blue fraction, agn fraction and star formation indicators from a targeted analysis of abell 1691. *Monthly Notices of the Royal Astronomical Society*, 426(2):1632–1646, 2012.

Ghanasyam Rallapalli, Diane GO Saunders, Kentaro Yoshida, Anne Edwards, Carlos A Lugo, Steve Collin, Bernardo Clavijo, Manuel Corpas, David Swarbreck, Matthew Clark, et al. Cutting edge: Lessons from fraxinus, a crowd-sourced citizen science game in genomics. *Elife*, 4:e07460, 2015.

Kelly S Reyna and Dale Rollins. Retention and efficacy of citizen scientist volunteers of the texas quail index. In *National Quail Symposium Proceedings*, volume 8, page 69, 2017.

KR Ridgway. Long-term trend and decadal variability of the southward penetration of the east australian current. *Geophysical Research Letters*, 34 (13), 2007.

R.A. Rocca, G. Magoon, D.F. Reynolds, T. Krahn, V.O. Tilroe, P.M. Op den Velde Boots, and A.J. Grierson. Discovery of western european r1b1a2 y chromosome variants in 1000 genomes project data: An online community approach. *PLoS ONE*, 7(7), jul 24 2012. ISSN 1932-6203.

R.G. Rocha, B.P. Vieira, V. Rodrigues, and C. Fonseca. Public engagement offers insights on the eurasian red squirrel distribution. *European Journal of Wildlife Research*, 63(6), December 2017. ISSN 1612-4642.

Katrin Ronnenberg, Egbert Strauß, and Ursula Siebert. Crop diversity loss as primary cause of grey partridge and common pheasant decline in lower saxony, germany. *BMC ecology*, 16(1):39, 2016.

Colleen Rothe-Groleau, Claudia M Rauter, and James D Fawcett. Morphological traits as indicators of sexual dimorphism in prairie rattlesnakes (crotalus viridis). 2018.

H.E. Roy and P.M.J. Brown. Ten years of invasion: Harmonia axyridis (pallas) (coleoptera: Coccinellidae) in britain. *Ecological Entomology*, 40(4):336–348, August 2015. ISSN 0307-6946.

Megan E. Schwamb, Chris J. Lintott, Debra A. Fischer, Matthew J. Giguere, Stuart Lynn, Arfon M. Smith, John M. Brewer, Michael Parrish, Kevin Schawinski, and Robert J. Simpson. Planet hunters: Assessing the kepler inventory of short-period planets. *ASTROPHYSICAL JOURNAL*, 754(2), aug 1 2012. ISSN 0004-637X.

David C Seburn and Erin Mallon. Has the eastern red-backed salamander (plethodon cinereus) declined in ontario? *The Canadian Field-Naturalist*, 131(2): 115–119, 2017.

Lior Shamir. Asymmetry between galaxies with clockwise handedness and counterclockwise handedness. *ASTROPHYSICAL JOURNAL*, 823(1), may 20 2016. ISSN 0004-637X.

F.M. Shilling and D.P. Waetjen. Wildlife-vehicle collision hotspots at us highway extents: Scale and data source effects. *Nature Conservation*, 11 (11, SI):41–60, 2015. ISSN 1314-6947. IENE International Conference on Ecology and Transportation, Malmo Univ, Malmo, SWEDEN, SEP 16-19, 2014.

RJ Simpson, MS Povich, S Kendrew, CJ Lintott, E Bressert, K Arvidsson, C Cyganowski, S Maddison, Kevin Schawinski, R Sherman, et al. The milky way project first data release: a bubblier galactic disc. *Monthly Notices of the Royal Astronomical Society*, 424(4):2442–2460, August 2012. ISSN 0035-8711.

A. M. Smith, S. Lynn, M. Sullivan, C. J. Lintott, P. E. Nugent, J. Botyanszki, M. Kasliwal, R. Quimby, S. P. Bamford, L. F. Fortson, K. Schawinski, I. Hook, S. Blake, P. Podsiadlowski, J. Joensson, A. Gal-Yam, I. Arcavi, D. A. Howell, J. S. Bloom, J. Jacobsen, S. R. Kulkarni, N. M. Law, E. O. Ofek, and R. Walters. Galaxy zoo supernovae star. *MONTHLY NOTICES OF THE ROYAL ASTRONOMICAL SOCIETY*, 412(2):1309–1319, April 2011. ISSN 0035-8711.

Bernardo Cervantes Sodi and Osbaldo Sánchez García. Do low surface brightness galaxies host stellar bars? *The Astrophysical Journal*, 847(1):37, 2017.

Gordon O SOHOLT. Goldstone apple valley radio telescope. In *2010 GSA Denver Annual Meeting*, 2010.

E. Solano, C. Rodrigo, R. Pulido, and B. Carry. Precovery of near-earth asteroids by a citizen-science project of the spanish virtual observatory. *Astronomische Nachrichten*, 335(2):142–149, February 2014. ISSN 0004-6337.

Peter Soroye, Najeeba Ahmed, and Jeremy T Kerr. Opportunistic citizen science data transform understanding of species distributions, phenology, and diversity gradients for global change research. *Global Change Biology*, 2018.

David John Stanton, Michael Robertson Leven, and Tommy Chung Hong Hui. Distribution of nanhaipotamon hongkongense (shen, 1940)(crustacea: Brachyura: Potamidae), a freshwater crab endemic to hong kong. *Journal of Threatened Taxa*, 10(1):11156–11165, 2018.

David A Steen, Dirk J Stevenson, Jeffrey C Beane, John D Willson, Matthew J Aresco, James C Godwin, Sean P Graham, Lora L Smith, Jennifer M Howze, D Craig Rudolph, et al. Terrestrial movements of the red-bellied mudsnake (farancia abacura) and rainbow snake (f. erytrogramma). *Herpetological Review 44 (2): 208-213*, 44(2): 208–213, 2013.

Catherine Sun, Angela K Fuller, and Jeremy E Hurst. Citizen science data enhance spatio-temporal extent and resolution of animal population studies. *bioRxiv*, page 352708, 2018.

Lee A Sutcliffe. Insights from project feederwatch: Changes in the abundance and occurrence of birds in new hampshire over the past 24 years. 2014.

Bruno Teixeira, Andre Hirsch, Vinicius D. L. R. Goulart, Luiza Passos, Camila P. Teixeira, Philip James, and Robert Young. Good neighbours: distribution of black-tufted marmoset (callithrix penicillata) in an urban environment. *WILDLIFE RESEARCH*, 42(7): 579–589, 2015. ISSN 1035-3712.

Jolyon Troscianko, Jared Wilson-Aggarwal, David Griffiths, Claire N Spottiswoode, and Martin Stevens. Relative advantages of dichromatic and trichromatic color vision in camouflage breaking. *Behavioral Ecology*, 28(2): 556–564, mar-apr 2017. ISSN 1045-2249.

Yves Turcotte, Jean-Francois Lamarre, and Joel Bety. Annual and seasonal variation in shorebird abundance in the st. lawrence river estuary during fall migration. *The Canadian Field-Naturalist*, 131(3):203–214, 2018.

B. Villarroel, I. Imaz, and J. Bergstedt. Our sky now and then: Searches for lost stars and impossible effects as probes of advanced extraterrestrial civilizations. *Astronomical Journal*, 152(3), September 2016. ISSN 0004-6256.

V Vinogradoff, S Bernard, C Le Guillou, and L Remusat. Evolution of interstellar organic compounds under asteroidal hydrothermal conditions. *Icarus*, 305: 358–370, 2018.

Constance E. Walker, Robert T. Sparks, and Stephen M. Pompea. Building on the international year of astronomy: The dark skies awareness program. In J Barnes, DA Smith, MG Gibbs, and JG Manning, editors, *SCIENCE EDUCATION AND OUTREACH: FORGING A PATH TO THE FUTURE*, volume 431 of *Astronomical Society of the Pacific Conference Series*, pages 103–109. Spitzer Sci Ctr; NASA Lunar Sci Inst; NASA Sci Miss Directorate; NRAO; California Space Grant Consortium; Northrop Grumman; Sky Skan; Univ Chicago Press; Ball Aerospace Technol; Capitol Coll; Astronom Assoc No California; Explore Sci; MWT Assoc Inc; Seiler Instruments; Astron Soc Pacific, 2010. ISBN 978-1-58381-742-1. 121st Annual Meeting of the Astronomical-Society-of-the-Pacific, Millbrae, CA, SEP 12-16, 2009.

N. Warning and D. Leatherman. Observations of canyon wren nestling diet during prey delivery. *Natural Areas Journal*, 36(2):181–186, April 2016. ISSN 0885-8608.

K.W. Willett, C.J. Lintott, S.P. Bamford, K.L. Masters, B.D. Simmons, K.R.V. Casteels, E.M. Edmondson, L.F. Fortson, S. Kaviraj, W.C. Keel, T. Melvin, R.C. Nichol, M. Jordan Raddick, K. Schawinski, R.J. Simpson, R.A. Skibba, A.M. Smith, and D. Thomas. Galaxy zoo 2: Detailed morphological classifications for 304 122 galaxies from the sloan digital sky survey. *Monthly Notices of the Royal Astronomical Society*, 435(4):2835–2860, November 2013. ISSN 0035-8711.

Russell E Winter. Hunting behaviors and foraging success of winter irruptive snowy owls in new york. 2016.

D.A. Woller and J.G. Hill. Melanoplus foxi hebard, 1923 (orthoptera: Acrididae: Melanoplinae): Rediscovered after almost 60 years using historical field notes connected to curated specimens. *Transactions of the American Entomological Society*, 141(3):545–574, December 2015. ISSN 0002-8320.

Jin-Long Xu, Di Li, Chuan-Peng Zhang, Xiao-Lan Liu, Jun-Jie Wang, Chang-Chun Ning, and Bing-Gang Ju. Gas kinematics and star formation in the filamentary irdc g34. 43+ 0.24. *The*

Astrophysical Journal, 819(2):117, 2016.

Jin-Long Xu, Ye Xu, Chuan-Peng Zhang, Xiao-Lan Liu, Naiping Yu, Chang-Chun Ning, and Bing-Gang Ju. Gas kinematics and star formation in the filamentary molecular cloud g47. 06+ 0.26. *Astronomy & Astrophysics*, 609:A43, 2018.

C-P Zhang, J-J Wang, and J-L Xu. Molecular clumps and star formation associated with the infrared dust bubble n131. *Astronomy & Astrophysics*, 550: A117, 2013.

A.11 Search or Result Errors

Incomplete records, about which further information could not be found, or which appear to result from errors in the search databases:

Alexis Ann Acohido and Manuel Paredes. 26 internships.

D Armitage, R Plummer, F Berkes, RI Arthur, AT Charles, IJ Davidson-Hunt, AP Diduck, N Doubleday, DS Johnson, and M Marschke. Aiken, lr 2002. attitudes and related psychosocial constructs: Theories, assess-ment, and research. thousand oaks, ca: Sage publications. aiken, lr, and dr aiken. 1969. recent research on attitudes concerning sci-ence. science education 53: 295–305. aikenhead, g., and a. ryan. 1992. the development of a new instrument: Views. *Citizen Science: Public Participation in Environmental Research*, 7:241, 2012.

EVAN A ESKEW and MICHAEL E DORCAS. Madeleine m. kern, jacquelyn c. guzy, steven j. price", stephanie d, hunt. *Journal of the North Carolina Academy of Science*, 128(3/4):81–87, 2012.

Washington Sea Grant. Citizen science.

S Griffiths et al. metry through crowd-sourcing experiences', internet archaeology 40. *Internet Archaeology*, 2015.

CJ Hansen. Epsc.

Yurong He and Andrea Wiggins. Special issue-(2016) volume 11 issue 15.

Digitally Invisible and DGL Team. An initiative of the johnson shoyama graduate school of public policy, university of regina.

John L King and Michael D Cohen. Radaphat chongthammakun.

Donald MacGregor, Marietta Baba, Aude Oliva, Anne Collins, Walt Scacchi McLaughlin, Brian Scassellati, Philip Rubin, Robert M Mason, James R Spohrer, Bruce Tonn, et al. Mark lundstrom and h.-s. philip wong.

Marcia O'Brien. Lara primary school. 2013.

Gregory MP O'Hare. Michael j. o'grady, conor muldoon, dominic carr, jie wan, barnard kroon.

Stepped Ramped. Citizen science results.

Holli Riebeek, Tamara Shapiro Ledley, David Herring, and Steven Lloyd. James acker.

Ph D Michael Stoltze. Program headquarters type data collection continent.

II Teil and Glossar zu verschiedenen Begriffen. Teil i citizen science praxis.

A.12 Social Media Citizen Science

Articles involving extracting Citizen Science data from social media channels:

N. Augar and M. Fluker. Developing social media for community based environmental monitoring. 2014.

Vijay Barve. Discovering and developing primary biodiversity data from social networking sites: A novel approach. *ECOLOGICAL INFORMATICS*, 24:194–199, November 2014. ISSN 1574-9541.

V. Boddula, A. Joshi, L. Ramaswamy, and D. Mishra. Harnessing social media for environmental sustainability: A measurement study on harmful algal blooms. pages 176–183, 2016.

N.A. Case, E.A. Macdonald, M. Heavner, A.H. Tapia, and N. Lalone. Mapping auroral activity with twitter. *Geophysical Research Letters*, 42(10): 3668–3672, 2015.

Stefan Daume. Mining twitter to monitor invasive alien species - an analytical framework and sample information topologies. *ECOLOGICAL INFORMATICS*, 31:70–82, January 2016. ISSN 1574-9541.

Stefan Daume and Kai Fueldner. '`forest tweets" - informal digital coverage of the oak processionary moth or why foresters should care about social media. *ALLGEMEINE FORST UND JAGDZEITUNG*, 187(9-10):185–197, 2016. ISSN 0002-5852.

D.-P. Deng, T.-R. Chuang, K.-T. Shao, G.-S. Mai, T.-E. Lin, R. Lemmens, C.-H. Hsu, H.-H. Lin, and M.-J. Kraak. Using social media for collaborative species identification and occurrence: Issues, methods, and tools. pages 22–29, 2012.

Paul S EARLE. How twitter can improve earthquake response. In *2010 GSA Denver Annual Meeting*, 2010.

Danielle LeMieux. Social and new media in space projects and networks: Preliminary findings of mixed-methods ethnographic research in australia, the us, and online. 2017.

Q.W. Lewis and E. Park. Volunteered geographic videos in physical geography: Data mining from youtube. *Annals of the American Association of Geographers*, 108(1):52–70, 2018. ISSN 2469-4452.

N. Michelsen, H. Dirks, S. Schulz, S. Kempe, M. Al-Saud, and C. SchÃ¯Â¿Â½th. Youtube as a crowd-generated water level archive. *Science of the Total Environment*, 568:189–195, 2016.

Emiliano Mori, Pietro Di Bari, and Marco Coraglia. Interference between roe deer and northern chamois in the italian alps: are facebook groups effective data sources? *ETHOLOGY ECOLOGY & EVOLUTION*, 30(3): 277–284, 2018. ISSN 0394-9370.

Arno Scharl, Michael Föls, David Herring, Lara Piccolo, Miriam Fernandez, and Harith Alani. Application design and engagement strategy of a game with a purpose for climate change awareness. In *International Conference on Internet Science*, pages 97–104. Springer, 2016.

Arno Scharl, Michael Fols, and David D Herring. Climate challenge–raising collective awareness in the tradition of games with a purpose. 2015.

Arno Scharl, Marta Sabou, and Michael Föls. Climate quiz: a web application for eliciting and validating knowledge from social networks. In *Proceedings of the 18th Brazilian symposium on Multimedia and the web*, pages 189–192. ACM, 2012.

Amy Stewart, Bianca Ambrose-Oji, and Jake Morris. Social media and forestry.

N. Suprayitno, R.P. Narakusumo, T. von Rintelen, L. Hendrich, and M. Balke. Taxonomy and biogeography without frontiers - whatsapp, facebook and smartphone digital photography let citizen scientists in more remote localities step out of the dark. *Biodiversity Data Journal*, 5, 2017.

H. Zheng, Y. Hong, D. Long, and H. Jing. Monitoring surface water quality using social media in the context of citizen science. *Hydrology and Earth System Sciences*, 21(2):949–961, 2017.

A.13 Specific Projects

Articles concerning specific Citizen Science projects:

Mimi Arandjelovic, Colleen R Stephens, Maureen S McCarthy, Paula Dieguez, Ammie K Kalan, Nuria Maldonado, Christophe Boesch, and Hjalmar S Kuehl. Chimp&see: an online citizen science platform for large-scale, remote video camera trap annotation of chimpanzee behaviour, demography and individual identification. *PeerJ Preprints*, 2016.

Maria Aristeidou, Eileen Scanlon, and Mike Sharples. Weather-it: evolution of an online community for citizen inquiry. In *Proceedings of the 15th International Conference on Knowledge Technologies and Data-driven Business*, volume 21-22-October-2015, page 13. ACM, 2015.

C. Artemi. From galaxyzoo to dark matter. 2010.

K Aye, Megan E Schwamb, Ganna Portyankina, Candice J Hansen, Adam McMaster, Grant RM Miller, Brian Carstensen, Christopher Snyder, Michael Parrish, Stuart Lynn, et al. Planet four: Probing seasonal winds on mars by mapping the southern polar co $ _2 $ jet deposits. *arXiv preprint arXiv:1803.10341*, 2018.

J. Bahamonde, A. Hevia, G. Font, J. Bustos-Jimenez, and C. Montero. Mining private information from public data: The transantiago case. *IEEE Pervasive Computing*, 13(2):37–43, 2014.

Julie K Banfield, OI Wong, Kylie W Willett, Ray P Norris, L Rudnick, Stanislav S Shabala, Brooke D Simmons, Chris Snyder, A Garon, Nick Seymour, et al. Radio galaxy zoo: host galaxies and radio morphologies derived from visual inspection. *Monthly Notices of the Royal Astronomical Society*, 453(3): 2326–2340, 2015.

Miri Barak and Shadi Asakle. Augmentedworld: Facilitating the creation of location-based questions. *COMPUTERS & EDUCATION*, 121:89–99, June 2018. ISSN 0360-1315.

L. Barnard, A.M. Portas, S.L. Gray, and R.G. Harrison. The national eclipseweather experiment: An assessment of citizen scientist. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 374(2077), sep 28 2016. ISSN 1364-503X.

L. Belbin and K.J. Williams. Towards a national bio-environmental data facility: experiences from the atlas of living australia. *International Journal of Geographical Information Science*, 30(1): 108–125, jan 2 2016. ISSN 1365-8816.

C. Bonacic, A. Neyem, and A. Vasquez. Live andes: Mobile-cloud shared workspace for citizen science and wildlife conservation. In *2015 IEEE 11TH INTERNATIONAL CONFERENCE ON E-SCIENCE*, Proceeding IEEE International Conference on e-Science (e-Science), pages 215–223. CPS Conf Publishing Serv; IEEE Comp Soc, 2015. ISBN 978-1-4673-9325-6. IEEE International Conference On eScience, Munich, GERMANY, AUG 31-SEP 04, 2015.

Caitlin Bonnar, Meg Campbell, Ryan Drapeau, Cynthia Bennett, and Alan Borning. Contributing during the commute: Why transit riders submit information about bus stops with stopinfo. 2015.

R. Bossu, F. Roussel, L. Fallou, M. LandÃ¨s, R. Steed, G. Mazet-Roux, A. Dupont, L. Frobert, and L. Petersen. Lastquake: From rapid information to global seismic risk reduction. *International Journal of Disaster Risk Reduction*, 28:32–42, June 2018. ISSN 2212-4209.

R Bugiolacchi, IA Crawford, and KH Joy. Moon zoo: a citizen science project. In *European Planetary Science Congress*, volume 8, 2013.

S.G. Caneva and B.F. van Oppen. Signet: A digital platform for hellenistic sealings and archives. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 10059 LNCS:222–231, 2016.

D.J. Cantrill. The australasian virtual herbarium: Tracking data usage and benefits for biological collections: Tracking. *Applications in Plant Sciences*, 6(2), February 2018. ISSN 2168-0450.

Matías Celasco, Juan Ignacio Yañez, Roberto Gamen, Alejandro Fernandez, Alicia Díaz, and Diego Torres. Galaxy conqueror: Astronomy, citizen science and gamification. In *Learning Objects and Technology (LACLO), Latin American Conference on*, pages 1–10. IEEE, 2016.

Kevin Crowston. Gravity spy: Humans, machines and the future of citizen science. In *Companion of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*, pages 163–166. ACM, 2017.

Heather R Cunningham, Charles A Davis, Christopher W Swarth, and Glenn D Therres. The maryland amphibian and reptile atlas: a volunteer-based distributional survey. *International Journal of Zoology*, 2012, 2012.

Travis Desell, Robert Bergman, Kyle Goehner, Ronald Marsh, Rebecca VanderClute, and Susan Ellis-Felege. Wildlife home: Combining crowd sourcing and volunteer computing to analyze avian nesting video. In *Wildlife Home: Combining Crowd Sourcing and Volunteer Computing to Analyze Avian Nesting Video*, pages 107–115. IEEE, 2013.

Jane Disney, Duncan Bailey, Anna Farrell, and Ashley Taylor. Next generation citizen science using anecdata. org. *Maine Policy Review*, 26(2):70–79, 2017.

Francis R Eanes, Janet M Silbernagel, David A Hart, Patrick Robinson, and Max Axler. Participatory mobile-and web-based tools for eliciting landscape knowledge and perspectives: introducing and evaluating the wisconsin geotools project. *Journal of Coastal Conservation*, 22(2):399–416, 2018.

Special Edition. Froglog.

Elizabeth R Ellwood, Paul Kimberly, Robert Guralnick, Paul Flemons, Kevin Love, Shari Ellis, Julie M Allen, Jason H Best, Richard Carter, Simon Chagnoux, et al. Worldwide engagement for digitizing biocollections (wedigbio): The biocollections community's citizen-science space on the calendar. *BioScience*, 68(2):112–124, February 2018. ISSN 0006-3568.

Jerry Alan Fails, Katherine G Herbert, Emily Hill, Christopher Loeschorn, Spencer Kordecki, David Dymko, Andrew DeStefano, and Zill Christian. Geotagger: a collaborative and participatory environmental inquiry system. In *Proceedings of the companion publication of the 17th ACM conference on Computer supported cooperative work & social computing*, pages 157–160. ACM, 2014.

Lucy Fortson, Karen Masters, Robert Nichol, Kirk D. Borne, Edward M. Edmondson, Chris Lintott, Jordan Raddick, Kevin Schawinski, and John Wallin. Galaxy zoo morphological classification and citizen science. In MJ Way, JD Scargle, KM Ali, and AN Srivastava, editors, *ADVANCES IN MACHINE LEARNING AND DATA MINING FOR ASTRONOMY*, Chapman & Hall-CRC Data Mining and Knowledge Discovery Series, pages 213–236. 2012a. ISBN 978-1-4398-4174-7; 978-1-4398-4173-0.

Lucy Fortson, Karen Masters, Robert Nichol, EM Edmondson, C Lintott, J Raddick, and J Wallin. Galaxy zoo. *Advances in machine learning and data mining for astronomy*, 2012:213–236, 2012b.

Francisco Jose Garcia Penalvo. Wyred project. *EDUCATION IN THE KNOWLEDGE SOCIETY*, 18(3):7–14, September 2017. ISSN 2444-8729.

P Gay, C Lehan, G Bracey, and N Gugiucci. Cosmoquest: A virtual facility for learning and doing science. In *European Planetary Science Congress 2012*, 2012.

Pamela L Gay and C Lintott. A zoo of online citizen science. In *American Astronomical Society Meeting Abstracts# 214*, volume 214, page 757, 2009.

W.T. Gee, O. Guyon, J. Walawender, N. Jovanovic, and L. Boucher. Project panoptes: A citizen-scientist exoplanet transit survey using commercial digital cameras. In CJ Evans, L Simard, and

H Takami, editors, *GROUND-BASED AND AIRBORNE INSTRUMENTATION FOR ASTRONOMY VI*, volume 9908 of *Proceedings of SPIE*. SPIE, 2016. ISBN 978-1-5106-0196-3. Conference on Ground-Based and Airborne Instrumentation for Astronomy VI, Edinburgh, SCOTLAND, JUN 26-30, 2016.

Rory Gibb, Oisin Mac Aodha, and E Jonesdescribe. Bat detective: citizen science for eco-acoustic biodiversity monitoring.

H. GoÃau, A. Affouard, S. Dufour, C. Vignau, P. Bonnet, V. Bakic, S. Selmi, D. BarthÃl Amy, A. Joly, J. Barbe, I. Yahiaoui, and N. Boujemaa. Plntnet mobile 2014: Android port and new features. pages 527–528, 2014.

Leslie Goldman. Neon citizen science: Planning and prototyping.

Nicole Gugliucci, Pamela Gay, Georgia Bracey, Irene Antonenko, Stuart Robbins, Britney E. Schmidt, Jennifer E. C. Scully, Cory Lehan, and Joseph Moore. Citizen science with cosmoquest: Science and strategies. In JG Manning, MK Hemenway, JB Jensen, and MG Gibbs, editors, *ENSURING STEM LITERACY: A NATIONAL CONFERENCE ON STEM EDUCATION AND PUBLIC OUTREACH*, volume 483 of *Astronomical Society of the Pacific Conference Series*, pages 237–240. EXOPLANET EXPLORAT PROGRAM; JPL; Stratospher Observ Infrared Astron; NRAO; E&S; Sky Skan; LOCKHEED MARTIN; sapl learn; Bell Aerosp & Technol Corp; AAS; EXPLORE SCI; SEILER INSTRUMENT; Astron Soc Pacif; San Jose State Univ, 2014. ISBN 978-1-58381-850-3. 125th ASP Annual Conference on Ensuring STEM Literacy: A National Conference on STEM Education and Public Outreach, San Jose State Univ, San Jose, CA, JUL 20-24, 2013.

Muki Haklay, Louise Francis, Charlene Jennett, and Laure Kloetzer. Community engagement around poor air quality in london: Citizen inquiry in a citizen science 'mapping for change' project. In *Citizen Inquiry*, pages 60–80. Routledge, 2017.

Andrew Hill, Robert Guralnick, Arfon Smith, Andrew Sallans, Rosemary Gillespie, Michael Denslow, Joyce Gross, Zack Murrell, Tim Conyers, Peter Oboyski, et al. The notes from nature tool for unlocking biodiversity records from museum records through citizen science. *ZooKeys*, 209(209):219, 2012. ISSN 1313-2989.

Catherine Hoffman, Caren B Cooper, Eric B Kennedy, Mahmud Farooque, and Darlene Cavalier. Scistarter 2.0: A digital platform to foster and study sustained engagement in citizen science. In *Analyzing the Role of Citizen Science in Modern Research*, Advances in Knowledge Acquisition Transfer and Management (AKATM) Book Series, pages 50–61. IGI Global, 2017. ISBN 978-1-5225-0963-9; 978-1-5225-0962-2.

C Jennett. Creativity in citizen cyberlab. 2015.

Charlene Jennett, DOMINIC Furniss, Ioanna Iacovides, Sarah Wiseman, Sandy JJ Gould, and Anna L Cox. Exploring citizen psych-science and the motivations of errordiary volunteers. *Human Computation*, 1(2):200–218, 2014.

S. Jirka, A. Remke, and A. BrÃ¶ring. Envirocar-crowd sourced traffic and environment data for sustainable mobility. volume 1322, 2013.

A. Joly, H. MÃller, H. GoÃau, H. Glotin, C. Spampinato, A. Rauber, P. Bonnet, W.-P. Vellinga, and B. Fisher. Lifeclef: Multimedia life species identification. volume 1222, pages 7–13, 2014.

Katherine Joy, Ian Crawford, Peter Grindrod, Chris Lintott, Steven Bamford, Arfon Smith, Anthony Cook, Irene Antonenko, Ryan Balfanz, Matt Balme, et al. Moon zoo: citizen science in lunar exploration. *Astronomy & Geophysics*, 52 (2):2–10, 2011.

Detlef Kanwischer, David Burger, and Thomas Nauss. Citizen science and digital geomedia: Implementing a biodiversity information system in cabo verde. In R Vogler, A Car, J Strobl, and G Griesebner, editors, *GI FORUM 2014: GEOSPATIAL INNOVATION FOR SOCIETY*, pages 299–308. Univ Salzburg, Interfaculty Dept Geoinformat Z GIS; Austrian Acad Sci, Commiss GISci; Univ Technol, Dept Sustainable Tourism & Reg Dev, 2014. ISBN 978-3-87907-545-4. Geoinformatics Forum, Salzburg, AUSTRIA, JUL 01-04, 2014.

A. D. Kapinska, I. Terentev, O. I. Wong, S. S. Shabala, H. Andernach, L. Rudnick, L. Storer, J. K. Banfield, K. W. Willett, F. de Gasperin, C. J. Lintott, A. R. Lopez-Sanchez, E. Middelberg, R. P. Norris, K. Schawinski, N. Seymour, and B. Simmons. Radio galaxy zoo: A search for hybrid morphology radio galaxies. *ASTRONOMICAL JOURNAL*, 154(6), December 2017. ISSN 0004-6256.

C.L. Kirkhope, R.L. Williams, C.L. Catlin-Groves, S.G. Rees, C. Montesanti, J. Jowers, H. Stubbs, J. Newberry, A.G. Hart, A.E. Goodenough, and R. Stafford. Social networking for biodiversity: The beeid project. pages 625–626, 2010.

Margaret Kosmala, Alycia Crall, Rebecca Cheng, Koen Hufkens, Sandra Henderson, and Andrew D Richardson. Season spotter: Using citizen science to validate and scale plant phenology from near-surface remote sensing. *Remote Sensing*, 8(9):726, September 2016. ISSN 2072-4292.

E. Kosmidis, P. Syropoulou, S. Tekes, P. Schneider, E. Spyromitros-Xioufis, M. Riga, P. Charitidis, A. Moumtzidou, S. Papadopoulos, S. Vrochidis, I. Kompatsiaris, I. Stavrakas, G. Hloupis, A. Loukidis, K. Kourtidis, A.K. Georgoulias, and G. Alexandri. Hackair: Towards raising awareness about air quality in europe by developing a collective online platform. *ISPRS International Journal of Geo-Information*, 7 (5), May 2018. ISSN 2220-9964.

M.J. Kuchner, S.M. Silverberg, A.S. Bans, S. Bhattacharjee, S.J. Kenyon, J.H. Debes, T. Currie, L. GarcÃ¯Â¿Â½a, D. Jung, C. Lintott, M. McElwain, D.L. Padgett, L.M. Rebull, J.P. Wisniewski, E. Nesvold, K. Schawinski, M.L. Thaller, C.A. Grady, J. Biggs, M. Bosch, T. Cernohous, H.A.D. Luca, M. Hyogo, L.L.W. Wah, A. Piipuu, and F. PiÃ¯Â¿Â½eiro. Disk detective: Discovery of new circumstellar disk candidates through citizen science. *Astrophysical Journal*, 830(2), oct 20 2016. ISSN 0004-637X.

D. Kwak, A. Kam, D. Becerra, Q. Zhou, A. Hops, E. Zarour, A. Kam, L. Sarmenta, M. Blanchette, and J. WaldispÃhl. Open-phylo: A customizable crowd-computing platform for multiple sequence alignment. *Genome Biology*, 14(10), 2013. ISSN 1465-6906.

Doris Jung-Lin Lee, Joanne Lo, Moonhyok Kim, and Eric Paulos. Crowdclass: Designing classification-based citizen science learning modules. In *Fourth AAAI Conference on Human Computation and Crowdsourcing*, 2016.

Gary B LEWIS. Earthtrek-a new global citizen science portal that helps scientists to engage the public in real research. In *2010 GSA Denver Annual Meeting*, 2010.

Shanshan Li, Dongwei Fan, Xing Gao, and Chenzhou Cui. Popular supernova project: A citizen science program based on amateur astronomical observations. In LG Chova, AL Martinez, and IC Torres, editors, *INTED2016: 10TH INTERNATIONAL TECHNOLOGY, EDUCATION AND DEVELOPMENT CONFERENCE*, INTED Proceedings, pages 7561–7569, 2016. ISBN 978-84-608-5617-7. 10th International Technology, Education and Development Conference (INTED), Valencia, SPAIN, MAR 07-09, 2016.

Roman Lukyanenko and Jeffrey Parsons. Increasing quality in citizen science by giving volunteers more data-creation freedom.

Kurt Luther, Scott Counts, Kristin B Stecher, Aaron Hoff, and Paul Johns. Pathfinder: an online collaboration environment for citizen scientists. In S Greenberg, SE Hudson, K Hinkley, M RingelMorris, and DR Olsen, editors, *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 239–248. ACM, 2009. ISBN 978-1-60558-246-7. 27th Annual CHI Conference on Human Factors in Computing Systems, Boston, MA, APR 04-09, 2009.

Alan Macnaughton. The prairie provinces butterfly atlas.

Michael J Madison. Commons at the intersection of peer production, citizen science, and big data: galaxy zoo. *Governing knowledge commons*, 209:215, 2014.

M.L. Maher and S. Abdellahi. Naturenet: An interaction design with a focus on crowdsourcing for community. *Communications in Computer and Information Science*, 618:50–55, 2016. ISSN 1865-0929. 18th International Conference on Human-Computer Interaction (HCI International), Toronto, CANADA, JUL 17-22, 2016.

Philip J Marshall, Aprajita Verma, Anupreeta More, Christopher P Davis, Surhud More, Amit Kapadia, Michael Parrish, Chris Snyder, Julianne Wilcox, Elisabeth Baeten, et al. Space warps–i. crowdsourcing the discovery of gravitational lenses. *Monthly Notices of the Royal Astronomical Society*, 455(2): 1171–1190, 2015.

S Martellos, F Attore, D Cesaroni, S Di Marco, D Petruzzella, O Spinelli, G Tallone, and A Mereu. Csmon-life: data from the people, data for the people. In *First ECSA Conference 2016-Citizen Science-Innovation in Open Science, Society and Policy*. DE, 2016.

Thomas Maskell, Clara Crivellaro, Robert Anderson, Tom Nappey, Vera Araújo-Soares, and Kyle Montague. Spokespeople: Exploring routes to action through citizen-generated data. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, page 405. ACM, 2018.

A.D. Mason, G. Michalakidis, and P.J. Krause. Tiger nation: Empowering citizen scientists. 2012.

K.L. Masters. Galaxy zoo: Outreach

and science hand in hand. *Proceedings of the International Astronomical Union*, 10 (H16):689–691, 2012.

Akira Masuda, Kun Zhang, and Takuya Maekawa. Sonic home: environmental sound collection game for human activity recognition. *Journal of Information Processing*, 24(2):203–210, 2016.

Patrick McAndrew, David Robinson, Michael Dodd, and Janice Ansine. Exploring citizen science and inquiry learning through ispotnature. org. In *Citizen Inquiry*, pages 83–103. Routledge, 2017.

Bryan J. H. Mendez, Brian Day, Pamela L. Gay, Suzanne H. Jacoby, M. Jordan Raddick, Constance E. Walker, and Stephen M. Pompea. The spectrum of citizen science projects in astronomy and space science. In J Barnes, DA Smith, MG Gibbs, and JG Manning, editors, *SCIENCE EDUCATION AND OUTREACH: FORGING A PATH TO THE FUTURE*, volume 431 of *Astronomical Society of the Pacific Conference Series*, pages 324–333. Spitzer Sci Ctr; NASA Lunar Sci Inst; NASA Sci Miss Directorate; NRAO; California Space Grant Consortium; Northrop Grumman; Sky Skan; Univ Chicago Press; Ball Aerospace Technol; Capitol Coll; Astronom Assoc No California; Explore Sci; MWT Assoc Inc; Seiler Instruments; Astron Soc Pacific, 2010. ISBN 978-1-58381-742-1. 121st Annual Meeting of the Astronomical-Society-of-the-Pacific, Millbrae, CA, SEP 12-16, 2009.

J.-O. Meynecke. Whale trails - a smart phone application for whale tracking. volume 1, pages 304–310, 2014.

A Minutolo, A Scialoja, G Zampetti, T Schäfer, B Kieslinger ZSI, Maria García González, Jorge Garcia Vidal, Jose M Barceló UPC, Sibylle Egger, Anna Ripoll, et al. Collective awareness platform for tropospheric ozone pollution.

Jules Morgan. Gaming for dementia research: a quest to save the brain. *The Lancet Neurology*, 15(13):1313, 2016.

J. Mueller, S. Doerfel, M. Becker, A. Hotho, and G. Stumme. Tag recommendations for sensorfolksonomies. volume 1066, 2013.

P.L. Nimis, A. Moro, F. Attorre, S. Martellos, and E. Chiancone. A digital flora of rome. *Plant Biosystems*, 150(3): 384–387, 2016. ISSN 1126-3504.

P.L. Nimis, R. Riccamboni, and S. Martellos. Identification keys on mobile devices: The dryades experience. *Plant Biosystems*, 146(4):783–788, 2012. ISSN 1126-3504.

Jill Nugent. Be a bat detective! *Science Scope*, 41(2):12, 2017.

Melissa Nursey-Bray, Robert Palmer, and Gretta Pecl. Spot, log, map: Assessing a marine virtual citizen science program against reed's best practice for stakeholder participation in environmental management. *Ocean & Coastal Management*, 151:1–9, jan 1 2018. ISSN 0964-5691.

C. Pennington, K. Freeborough, C. Dashwood, T. Dijkstra, and K. Lawrie. The national landslide database of great britain: Acquisition, communication and the role of social media. *Geomorphology*, 249(SI):44–51, nov 15 2015. ISSN 0169-555X.

L. Peruzzi, S. Bagella, R. Filigheddu, B. Pierini, M. Sini, F. Roma-Marzio, K.F. Caparelli, G. Bonari, G. Gestri, D. Dolci, A. Consagra, P. Sassu, M.C. Caria, G. Rivieccio, M. Marrosu, M. D'Antraccoli, G. Pacifico, V. Piu, and G. Bedini. The wikiplantbase project: the role of amateur botanists in building up large online floristic databases. *Flora Mediterranea*, 27:117–129, 2017.

Richard G Pinet. Naturewatch 2.0: Reconnecting canadians with nature in a digital age. In *EdMedia: World Conference on Educational Media and Technology*, pages 1603–1603. Association for the Advancement of Computing in Education (AACE), 2016.

J. Powell, G. Nash, and P. Bell. Geoexposures: Documenting temporary geological exposures in great britain through a citizen-science web site. *Proceedings of the Geologists' Association*, 124(4): 638–647, June 2013. ISSN 0016-7878.

S.J. Price and M.E. Dorcas. The carolina herp atlas: An online, citizen-science approach to document amphibian and reptile occurrences. *Herpetological Conservation and Biology*, 6(2): 287–296, August 2011. ISSN 1931-7603.

Kathleen L Prudic, Kent P McFarland, Jeffrey C Oliver, Rebecca A Hutchinson, Elizabeth C Long, Jeremy T Kerr, and Maxim Larrivée. ebutterfly: leveraging massive online citizen science for butterfly conservation. *Insects*, 8(2):53, June 2017. ISSN 2075-4450.

V. Ranner. Uisilk - towards interfacing the body. pages 13–17, 2013.

Henry W REGES, Nolan J DOESKEN, Noah NEWMAN, Zach SCHWALBE, and Julian TURNER. Cocorahs (the community collaborative rain, hail and snow network); on-line citizen science engaging the public in precipitation monitoring. *Abstracts with Programs–Geological Society of America, Denver, 2010*, 2010.

D. Riemann, R. Glaser, M. Kahle, and S. Vogt. The cre tambora.org – new data and tools for collaborative research in climate and environmental history. *Geoscience Data Journal*, 2(2):63–77, November 2015. ISSN 2049-6060.

David Robinson, Michael Dodd, and Janice Ansine. Exploring citizen science and inquiry learning through ispotnature. org. In *Citizen Inquiry*, pages 101–121. Routledge, 2017a.

Kelly L Robinson, Jessica Y Luo, Su Sponaugle, Cedric Guigand, and Robert K Cowen. A tale of two crowds: Public engagement in plankton classification. *Frontiers in Marine Science*, 4:82, 2017b.

M. Roderick and J. Gross. The amphibiaweb app and use of mobile devices in research and outreach. *Herpetology Notes*, 7(0):109–113, 2014.

Rachel J. Ross. Las cumbres observatory global telescope network: Keeping citizen scientists in the dark. In JB Jensen, JG Manning, MG Gibbs, and D Daon, editors, *CONNECTING PEOPLE TO SCIENCE: A NATIONAL CONFERENCE ON SCIENCE EDUCATION AND PUBLIC OUTREACH*, volume 457 of *Astronomical Society of the Pacific Conference Series*, pages 339–342. Space Telescope Sci Inst; NASA Lunar Sci Inst; NASAs Explorat Program; Infrared Process & Anal Ctr; NASAs Herschel Sci Ctr; Spitzer Sci Ctr; Stratospher Observ Infrared Astron; NASAs Chandra XRay Observ; Univ Chicago Press; Natl Radio Astron Observ; Ball Aerospace; Capitol Coll; Sky Skan; Univ Wyoming CAPER Team; Amer Astron Soc; AAS Educ Off; Solar Dynam Observ; Seiler Instrument; Explore Sci; Pratt St Ale House; MWT Assoc Inc; Amer Elements; Charlesbridge; Celestron, 2012. ISBN 978-1-58381-796-4. 123rd Annual Meeting of the Astronomical-Society-of-the-Pacific, Amer Geophys Union, Baltimore, MD, JUL 30-AUG 03, 2011.

Robert M ROSS. Fossil finders: Invertebrate paleontological research carried out in partnership with 5th to 9th grade teachers and their students. In *2010 GSA Denver Annual Meeting*, 2010.

Dana Rotman, Kezia Procita, Derek Hansen, Cynthia Sims Parr, and Jennifer Preece. Supporting content curation communities: The case of the encyclopedia of life. *JOURNAL OF THE AMERICAN SOCIETY FOR INFORMATION SCIENCE AND TECHNOLOGY*, 63(6):1092–1107, June 2012. ISSN 1532-2882.

V. Runnel, M. Peterson, and A. Zirk. "my naturesound" - nature observations with sound recordings. *Biodiversity Data Journal*, 5, 2017.

J.J.W.H. SÃ̧rensen, M.K. Pedersen, M. Munch, P. Haikka, J.H. Jensen, T. Planke, M.G. Andreasen, M. Gajdacz, K. MÃ̧lmer, A. Lieberoth, and J.F. Sherson. Exploring the quantum speed limit with computer games. *Nature*, 532(7598): 210–213, apr 14 2016. ISSN 0028-0836.

K. Schrier. What's in a name? naming games that solve real-world problems. In A Canossa, C Harteveld, J Zhu, M Sicart, and S Deterding, editors, *PROCEEDINGS OF THE 12TH INTERNATIONAL CONFERENCE ON THE FOUNDATIONS OF DIGITAL GAMES (FDG'17)*, volume Part F130151, 2017. ISBN 978-1-4503-5319-9. 12th International Conference on the Foundations of Digital Games (FDG), Cape Cod, MA, AUG 14-17, 2017.

Megan E Schwamb. Highlights from citizen science in astronomy and planetary science. 2015.

Megan E Schwamb, Klaus-Michael Aye, Ganna Portyankina, Candice J Hansen, Campbell Allen, Sarah Allen, Fred J Calef III, Simone Duca, Adam McMaster, and Grant RM Miller. Planet four: Terrains–discovery of araneiforms outside of the south polar layered deposits. *Icarus*, 308:148–187, jul 1 2018. ISSN 0019-1035. 6th International Conference on Mars Polar Science and Exploration, Reykjavik, ICELAND, SEP, 2016.

Megan E Schwamb, Debra Fischer, Tabetha S Boyajian, Matthew J Giguere, Sascha Ishikawa, Chris Lintott, Stuart Lynn, Joseph Schmitt, Chris Snyder, Ji Wang, et al. Planet hunters 2 in the k2 era. In *American Astronomical Society Meeting Abstracts*, volume 225, 2015.

Megan E Schwamb, C Lintott, D Fischer, AM Smith, TS Boyajian, JM Brewer, MJ Giguere, S Lynn, M Parrish, K Schawinski, et al. Planet hunters: Kepler by eye. In *American Astronomical Society Meeting Abstracts# 223*, volume 223, 2014.

Megan E Schwamb, CJ Lintott, D Fischer, JM Brewer, MJ Giguere, S Lynn, M Parrish, Kevin Schawinski, R Simpson, AM Smith, et al. Planet hunters in the kepler extended mission. In *American Astronomical Society Meeting Abstracts# 221*, volume 221, 2013.

Catherine M Scott. Using citizen science to engage preservice elementary educators in scientific fieldwork. *Journal of College Science Teaching*, 46(2), 2016.

M. Sharples, M. Aristeidou,

E. Villasclaras-FernÃndez, C. Herodotou, and E. Scanlon. The sense-it app: A smartphone sensor toolkit for citizen inquiry learning. *International Journal of Mobile and Blended Learning*, 9(2): 16–38, apr-jun 2017. ISSN 1941-8647.

Sally Shuttleworth. Old weather: citizen scientists in the 19th and 21st centuries. *Science Museum Group Journal*, 3(3), 2016.

Jonathan Silvertown, Martin Harvey, Richard Greenwood, Mike Dodd, Jon Rosewell, Tony Rebelo, Janice Ansine, and Kevin McConway. Crowdsourcing the identification of organisms: A case-study of ispot. *ZooKeys*, (480):125, 2015.

V.S. Smith, S.D. Rycroft, K.T. Harman, B. Scott, and D. Roberts. Scratchpads: A data-publishing framework to build, share and manage information on the diversity of life. *BMC Bioinformatics*, 10 (SUPPL.14):S6, 2009. ISSN 1471-2105.

K. Sparks, A. Klippel, J.O. WallgrÃn, and D. Mark. Citizen science land cover classification based on ground and aerial imagery. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 9368:289–305, 2015. ISSN 0302-9743. 12th International Conference on Spatial Information Theory (COSIT), Santa Fe, NM, OCT 12-16, 2015.

James Sprinks, Jessica Wardlaw, and Gary Priestnall. Marscape: Exploring the martian landscape through parm (projected augmented relief model). In *EGU General Assembly Conference Abstracts*, volume 19, page 16620, 2017.

James Christopher Sprinks, Jessica Wardlaw, Robert Houghton, Steven Bamford, and Stuart Marsh. Mars in motion: An online citizen science platform looking for changes on the surface of mars. In *AAS/Division for Planetary Sciences Meeting Abstracts*, volume 48, 2016.

Esperanza Stancioff, Medea Steinman, Beth Bisson, and Abraham J Miller. Signs of the seasons: A maine phenology project.

J. Steinke and J. van Etten. Gamification of farmer-participatory priority setting in plant breeding: Design and validation of 'agroduos'. *Journal of Crop Improvement*, 31(3):356–378, 2017.

Alexandra Swanson, Margaret Kosmala, Chris Lintott, and Craig Packer. A generalized approach for producing, quantifying, and validating citizen science data from wildlife images. *Conservation Biology*, 30(3):520–531, 2016.

Alexandra Swanson, Margaret Kosmala, Chris Lintott, Robert Simpson, Arfon Smith, and Craig Packer. Snapshot serengeti, high-frequency annotated camera trap images of 40 mammalian species in an african savanna. *Scientific data*, 2: 150026, 2015.

H. Tangmunarunkit, C.K. Hsieh, B. Longstaff, S. Nolen, J. Jenkins, C. Ketcham, J. Selsky, F. Alquaddoomi, D. George, J. Kang, Z. Khalapyan, J. Ooms, N. Ramanathan, and D. Estrin. Ohmage: A general and extensible end-to-end participatory sensing platform. *ACM Transactions on Intelligent Systems and Technology*, 6(3), 2015.

A. Tapia, N. Lalone, E. MacDonald, N. Case, M. Hall, and M. Heavner. Aurorasaurus: Citizen science, early warning systems and space weather. volume WS-14-20, pages 30–32, 2014.

R. Tapia-McClung. A digital platform to support citizen-government interactions from volunteered geographic information and crowdsourcing in mexico city. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 9788:167–182, 2016. ISSN 0302-9743. 16th International Conference on Computational Science and Its Applications (ICCSA), Beijing, PEOPLES R CHINA, JUL 04-07, 2016.

Glenn D Therres, Charles A Davis, and Christopher W Swarth. Grid-based amphibian and reptile atlas using active searching: Apilot project. *Maryland Naturalist*, 53(1):33–51.

B.M. Thiers and R.E. Halling. The macrofungi collection consortium:. *Applications in Plant Sciences*, 6(2), February 2018. ISSN 2168-0450.

Tyrone J Tolbert, Shinnosuke Nakayama, and Maurizio Porfiri. Tracking nemo: Help scientists understand zebrafish behavior. *Zebrafish*, 15(3): 310–313, June 2018. ISSN 1545-8547.

S. VÃtek. Open source database of images deimos: Extension for large-scale subjective image quality assessment. In AG Tescher, editor, *APPLICATIONS OF DIGITAL IMAGE PROCESSING XXXVII*, volume 9217 of *Proceedings of SPIE*. SPIE, 2014. ISBN 978-1-62841-244-4. Conference on Applications of Digital Image Processing XXXVII, San Diego, CA, AUG 18-21, 2014.

J. Vandenbroucke, S. Bravo, P. Karn, M. Meehan, M. Plewa, T. Ruggles, D. Schultz, J. Peacock, and A.L. Simons. Detecting particles with cell phones: the distributed electronic cosmic-ray observatory. volume 30-July-2015, 2015.

VIJAY CHARAN VENKATACHALAM. Land cover validation game. 2015.

Michelle Viotti, Cassie Bowman, Tim Harris, Marc Mercuri, and Brian Satorius. Nasa's 'be a martian'citizen science project. In *Presentation at the Annual Meeting of the Geological Society of America, Denver*, 2010.

Constance E WALKER. The globe at night campaign: Shedding light on light pollution. In *2010 GSA Denver Annual Meeting*, 2010.

K. Wallace, S. Snedigar, and C. Cameron. 'is ash falling?', an online ashfall reporting tool in support of improved ashfall warnings and investigations of ashfall processes. *Journal of Applied Volcanology*, 4(1), 2015.

SHG Walter, J-P Muller, P Sidiropoulos, Y Tao, K Gwinner, ARD Putri, J-R Kim, R Steikert, S van Gasselt, G Michael, et al. The web-based interactive mars analysis and research system for the imars project. *Earth and Space Science*, 2018.

Y. Wang, N. Kaplan, G. Newman, and R. Scarpino. Citsci.org: A new model for managing, documenting, and sharing citizen science data. *PLoS Biology*, 13 (10):1–5, 2015.

Sky Watch. Nature watch, 2014.

Ruth West, Todd Margolis, Jarlath O'Neil-Dunne, and Eitan Mendelowitz. Metatree: augmented reality narrative explorations of urban forests. In *The Engineering Reality of Virtual Reality 2012*, volume 8289, page 82890G. International Society for Optics and Photonics, 2012.

D.L. White, R.P. Pargas, A.T. Chow, J. Chong, M. Cook, and I. Tak. The vanishing firefly project: Engaging citizen scientists with a mobile technology and real-time reporting framework. pages 85–92, 2014.

L.C. Wicks, G.S. Cairns, J. Melnyk, S. Bryce, R.R. Duncan, and P.A. Dalgarno. Enlightenment: High resolution smartphone microscopy as an educational and public engagement platform [version 1; referees: 2 approved]. *Wellcome Open Research*, 2, 2017.

C.G. Willis, E. Law, A.C. Williams, B.F. Franzone, R. Bernardos, L. Bruno, C. Hopkins, C. Schorn, E. Weber, D.S. Park, and C.C. Davis. Crowdcurio: an online crowdsourcing platform to facilitate climate change studies using herbarium specimens. *New Phytologist*, 215(1):479–488, July 2017. ISSN 0028-646X.

Lucas A Wilson and David P Anderson. Vast: Integrating volunteer computing into existing cyberinfrastructure. 2017.

J.P. Worthington, J. Silvertown, L. Cook, R. Cameron, M. Dodd, R.M. Greenwood, K. Mcconway, and P. Skelton. Evolution megalab: A case study in citizen science methods. *Methods in Ecology and Evolution*, 3(2): 303–309, April 2012. ISSN 2041-210X.

S. Wylie and L. Albright. Wellwatch: Reflections on designing digital media for multi-sited para-ethnography. *Journal of Political Ecology*, 21(1):320–348, 2014.

Poonam Yadav, Ioannis Charalampidis, Jeremy Cohen, John Darlington, and Francois Grey. A collaborative citizen science platform to bring together scientists, volunteers, and game players. *arXiv preprint arXiv:1707.09566*, 2017a.

Poonam Yadav, Ioannis Charalampidis, Jeremy Cohen, John Darlington, and Francois Grey. A collaborative citizen science platform for real-time volunteer computing and games. *IEEE Transactions on Computational Social Systems*, 5 (1):9–19, March 2018. ISSN 2329-924X.

Poonam Yadav, Jeremy Cohen, and John Darlington. Citizengrid: An online middleware for crowdsourcing scientific research. *arXiv preprint arXiv:1707.09489*, 2017b.

A.14 Specific Technologies/Tools

Articles offering information about specific hardware and tools:

B. Alberton, R.D.S. Torres, L.F. Cancian, B.D. Borges, J. Almeida, G.C. Mariano, J.D. Santos, and L.P.C. Morellato. Introducing digital cameras to monitor plant phenology in the tropics: applications for conservation. *Perspectives in Ecology and Conservation*, 15(2): 82–90, apr-jun 2017. ISSN 2530-0644.

Tom August, Martin Harvey, Paula Lightfoot, David Kilbey, Timos Papadopoulos, and Paul Jepson. Emerging technologies for biological recording. *Biological Journal of the Linnean Society*, 115(3):731–749, 2015.

R. Bossu, R. Steed, G. Mazet-Roux, C. Etivant, and F. Roussel. The emsc tools used to detect and diagnose the impact of global earthquakes from direct and indirect eyewitnesses' contributions. volume 2015-January, 2015.

E.S. Bradley and K.C. Clarke. Outdoor webcams as geospatial sensor networks: Challenges, issues and opportunities. *Cartography and Geographic Information Science*, 38(1):3–19, January 2011. ISSN 1523-0406.

TD Chalasani, M Maddah, and R Lukyanenko. Powering citizen science with inference clouds.

F. Danielsen, M. Enghoff, E. Magnussen, T. Mustonen, A. Degteva, K.K. Hansen, N. Levermann, S.D. Mathiesen, and Ã˜. Slettemark. *Citizen science tools for engaging local stakeholders and promoting local and traditional knowledge in landscape stewardship*. 2017.

M.A. Dunlap, A.H.T. Tang, and S. Greenberg. Applying geocaching principles to site-based citizen science and eliciting reactions via a technology probe. *Personal and Ubiquitous Computing*, 19(5-6):897–913, August 2015. ISSN 1617-4909.

B.C. Ebner, D. Starrs, D.L. Morgan, C.J. Fulton, J.A. Donaldson, J. Sean Doody, S. Cousins, M. Kennard, G. Butler, Z. Tonkin, S. Beatty, B. Broadhurst, R. Clear, M. Lintermans, and C.S. Fletcher. Emergence of field-based underwater video for understanding the ecology of freshwater fishes and crustaceans in australia. *Journal of the Royal Society of Western Australia*, 97(2):287–296, 2014.

A. Farina, P. James, C. Bobryk, N. Pieretti, E. Lattanzi, and J. McWilliam. Low cost (audio) recording (lcr) for advancing soundscape ecology towards the conservation of sonic complexity and biodiversity in natural and urban landscapes. *Urban Ecosystems*, 17(4):923–944, December 2014. ISSN 1083-8155.

E.J. Farnsworth, M. Chu, W.J. Kress, A.K. Neill, J.H. Best, J. Pickering, R.D. Stevenson, G.W. Courtney, J.K. Vandyk, and A.M. Ellison. Next-generation field guides. *BioScience*, 63(11):891–899, November 2013. ISSN 0006-3568.

Rebecca Ruth McIntosh, Ross Holmberg, and Peter Dann. Looking without landing–using remote piloted aircraft to monitor fur seal populations without disturbance. *Frontiers in Marine Science*, 5(JUN):202, 2018.

A.L. Nguy-Robertson, E.M. Brinley Buckley, A.S. Suyker, and T.N. Awada. Determining factors that impact the calibration of consumer-grade digital cameras used for vegetation analysis. *International Journal of Remote Sensing*, 37(14):3365–3383, July 2016. ISSN 0143-1161.

A.A. Royem, C.K. Mui, D.R. Fuka, and M.T. Walter. Technical note: Proposing a low-tech, affordable, accurate stream stage monitoring system. *Transactions of the ASABE*, 55(6):2237–2242, 2012.

Douglas Sheil, Badru Mugerwa, and Eric H Fegraus. African golden cats, citizen science, and serendipity: tapping the camera trap revolution. *South African journal of wildlife research*, 43(1): 74–78, April 2013. ISSN 0379-4369.

R. D. Stevenson, William A. Haber, and Robert A. Morris. Electronic field guides and user communities in the eco-informatics revolution. *Conservation Ecology*, 7(1), June 2003. ISSN 11955449.

W. Teng and A. Albayrak. Framework for processing citizen science data for applications to nasa earth science missions. 2017.

Audrey Verma, René van der Wal, and Anke Fischer. Imagining wildlife: New technologies and animal censuses, maps and museums. *Geoforum*, 75:75–86, October 2016. ISSN 0016-7185.

K. Wendelsdorf and S. Shah. Empowered genome community: Leveraging a bioinformatics platform as a citizen-scientist collaboration tool. *Applied and Translational Genomics*, 6:7–10, September 2015. ISSN 2212-0661.

C Richard Ziegler, J Angus Webb, Susan B Norton, Andrew S Pullin, and Andreas H Melcher. Digital repository of associations between environmental variables: a new resource to facilitate knowledge synthesis. *Ecological Indicators*, 53:61–69, June 2015. ISSN 1470-160X.

A.15 Use Cases

Articles describing use cases of Citizen Science data (e.g., for machine learning):

Fraser Baker, Claire L Smith, and Gina Cavan. A combined approach to classifying land surface cover of urban domestic gardens using citizen science data and high resolution image analysis. *Remote Sensing*, 10(4):537, 2018.

Connor Bowley, Marshall Mattingly, Andrew Barnas, Susan Ellis-Felege, and Travis Desell. Detecting wildlife in unmanned aerial systems imagery using convolutional neural networks trained with an automated feedback loop. In *International Conference on Computational Science*, pages 69–82. Springer, 2018.

Connor Ryan Bowley. *Training Convolutional Neural Networks Using an Automated Feedback Loop to Estimate the Population of Avian Species*. The University of North Dakota, 2017.

Ricardo Contreras, Gabriel Salazar, M Angelica Pinninghoff, and Neil Nagar. Neural networks on galaxy pattern recognition. In *Pattern Recognition Systems (ICPRS 2017), 8th International Conference of*, pages 1–6. IET, 2017.

Joana Costa, Catarina Silva, Mário Antunes, and Bernardete Ribeiro. On using crowdsourcing and active learning to improve classification performance. In *Intelligent Systems Design and Applications (ISDA), 2011 11th International Conference on*, pages 469–474. IEEE, 2011.

Kevin Crowston, Carsten Østerlund, and Tae Kyoung Lee. Blending machine and human learning processes. In *Proceedings of the 50th Hawaii International Conference on System Sciences*, 2017.

D.-P. Deng, G.-S. Mai, T.-R. Chuang, R. Lemmens, and K.-T. Shao. Social web meets sensor web: From user-generated content to linked crowdsourced observation data. volume 1184, 2014.

K.J. Edwards and M.M. Gaber. Identifying uncertain galaxy morphologies using unsupervised learning. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 7895 LNAI(PART 2):146–157, 2013. ISSN 0302-9743. 12th International Conference on Artificial Intelligence and Soft Computing, ICAISC 2013, Zakopane, POLAND, JUN 09-13, 2013.

C.J. Ferster and N.C. Coops. Integrating volunteered smartphone data with multispectral remote sensing to estimate forest fuels. *International Journal of Digital Earth*, 9(2):171–196, feb 1 2016. ISSN 1753-8947.

S. Gray, R. Jordan, A. Crall, G. Newman, C. Hmelo-Silver, J. Huang, W. Novak, D. Mellor, T. Frensley, M. Prysby, and A. Singer. Combining participatory modelling and citizen science to support volunteer conservation action. *Biological Conservation*, 208 (SI):76–86, April 2017. ISSN 0006-3207.

P. Hartley, R. Flamary, N. Jackson, A. S. Tagore, and R. B. Metcalf. Support vector machine classification of strong gravitational lenses. *MONTHLY NOTICES OF THE ROYAL ASTRONOMICAL SOCIETY*, 471(3):3378–3397, November 2017. ISSN 0035-8711.

Yun Huang, John Zimmerman, Anthony Tomasic, and Aaron Steinfeld. Combining contribution interactions to increase coverage in mobile participatory sensing systems. In *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services*, pages 365–376. ACM, 2016.

Nuala Hunt, Michael J O'Grady, Conor Muldoon, Barnard Kroon, Tom Rowlands, Jie Wan, and Gregory MP O'Hare. Citizen science: A learning paradigm for the smart city? *IxD&A*, 27: 44–65, 2015.

Darrel Jenerette, Jun Wang, Mark Chandler, Julie Ripplinger, Sofia Koutzoukis, Cui Ge, Lorena Castro Garcia, Dion Kucera, and Xiong Liu. Resolving uncertainties in the urban air quality, climate, and vegetation nexus through citizen science, satellite imagery, and atmospheric modeling (281632). In *2017 Fall Meeting*, 2017.

R. KÃng, P. Saha, A. More, E. Baeten, J. Coles, C. Cornen, C. Macmillan, P. Marshall, S. More, J. Odermatt, A. Verma, and J.K. Wilcox. Gravitational lens modelling in a citizen science context. *Monthly Notices of the Royal Astronomical Society*, 447(3):2170–2180, mar 1 2015. ISSN 0035-8711.

Tarun Reddy Katapally, Jasmin Bhawra, Scott T Leatherdale, Leah Ferguson, Justin Longo, Daniel Rainham, Richard Larouche, and Nathaniel Osgood. The smart study, a mobile health and citizen science methodological platform for active living surveillance, integrated knowledge translation, and policy interventions: Longitudinal study. *JMIR public health and surveillance*, 4(1), 2018.

B.C. Kosar, E.A. MacDonald, N.A. Case, Y. Zhang, E.J. Mitchell, and R. Viereck. A case study comparing citizen science aurora data with global auroral boundaries derived from satellite imagery and empirical models. *Journal of Atmospheric and Solar-Terrestrial Physics*, 2018.

A.C. Kumar and S.M. Bhandarkar. A deep learning paradigm for detection of harmful algal blooms. In *2017 IEEE WINTER CONFERENCE ON APPLICATIONS OF COMPUTER VISION (WACV 2017)*, IEEE Winter Conference on Applications of Computer Vision, pages 743–751. IEEE; IEEE Comp Soc; IEEE Biometr Council, 2017. ISBN 978-1-5090-4822-9. 17th IEEE Winter Conference on Applications of Computer Vision (WACV), Santa Rosa, CA, MAR 24-31, 2017.

Evan Kuminski, Joe George, John Wallin, and Lior Shamir. Combining human and machine learning for morphological analysis of galaxy images. *Publications of the Astronomical Society of the Pacific*, 126(944):959–967, October 2014. ISSN 00046280, 15383873.

JC Martin. Contributions of citizen science to landscape democracy: potentials and challenges of current approaches. 2017.

L. Milanesi, M. Pilotti, and B. Bacchi. Using web-based observations to identify thresholds of a person's stability in a flow. *Water Resources Research*, 52 (10):7793–7805, October 2016. ISSN 0043-1397.

M. Molinier, C.A. LÃpez SÃnchez, T. Toivanen, I. Korpela, J.J. Corral-Rivas, R. Tergujeff, and T. HÃme. Relasphone-mobile and participative in situ forest biomass measurements supporting satellite image mapping. *Remote Sensing*, 8(10), October 2016. ISSN 2072-4292.

Hung Nguyen, Sarah J Maclagan, Tu Dinh Nguyen, Thin Nguyen, Paul Flemons, Kylie Andrews, Euan G Ritchie, and Dinh Phung. Animal recognition and identification with deep convolutional neural networks for automated wildlife monitoring. In *Data Science and Advanced Analytics (DSAA), 2017 IEEE International Conference on*, pages 40–49. IEEE, 2017.

Eduardo Palermo, Jeffrey Laut, Oded Nov, Paolo Cappa, and Maurizio Porfiri. A natural user interface to integrate citizen science and physical exercise. *PloS one*, 12(2):e0172587, feb 23 2017a. ISSN 1932-6203.

Eduardo Palermo, Jeffrey Laut, Oded Nov, Paolo Cappa, and Maurizio Porfiri. Spatial memory training in a citizen science context. *Computers in Human Behavior*, 73:38–46, August 2017b. ISSN 0747-5632.

Sarah Parcak, Gregory Mumford, and Chase Childs. Using open access satellite data alongside ground based remote sensing: An assessment, with case studies from egypt's delta. *Geosciences*, 7(4): 94, 2017.

Gary RANDOLPH. Using globe student data and classroom resources to aid the understanding of earth system science. In *2010 GSA Denver Annual Meeting*, 2010.

Isaac S Rosenthal, Jarrett EK Byrnes, Kyle C Cavanaugh, Tom W Bell, Briana Harder, Alison J Haupt, Andrew TW Rassweiler, Alejandro Pérez-Matus, Jorge Assis, Ali Swanson, et al. Floating forests: Quantitative validation of citizen science data generated from consensus classifications. *arXiv preprint arXiv:1801.08522*, 2018.

Samantha Rowbotham, Merryn McKinnon, Joan Leach, Rod Lamberts, and Penelope Hawe. Does citizen science have the capacity to transform population health science? *Critical Public Health*, pages 1–11, 2017.

S. Rowe and N. Alexander. Citizen science: Does it make sense for nutrition communication? *Nutrition Today*, 51(6): 301–304, 2016.

W.J. Scheirer, S.E. Anthony, K. Nakayama, and D.D. Cox. Perceptual annotation: Measuring human vision to improve computer vision. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 36(8):1679–1686, August 2014. ISSN 0162-8828.

DANIEL Schneider, JOSE LUIS Fernandez-Marquez, and JULIEN Da Costa. Data analytics in citizen cyberscience: Evaluating participant learning and engagement with analytics. *Hum. Comput*, 3:69–97, 2016.

A. Siddharthan, C. Lambin, A.-M. Robinson, N. Sharma, R. Comont, E. O'mahony, C. Mellish, and R. Van Der Wal. Crowdsourcing without a crowd: Reliable online species identification using bayesian models to minimize crowd size. *ACM Transactions on Intelligent Systems and Technology*, 7(4), July 2016. ISSN 2157-6904.

Zachary D Siegel, Naihui Zhou, Scott Zarecor, Nigel Lee, Darwin A Campbell, Carson M Andorf, Dan Nettleton, Carolyn J Lawrence-Dill, Baskar Ganapathysubramanian, Iddo Friedberg, et al. Crowdsourcing image analysis for plant phenomics to generate ground truth data for machine learning. *bioRxiv*, 2018.

J. Thomas, A. Noel-Storr, I. Marshall, B. Wallace, S. McDonald, C. Mavergames, P. Glasziou, I. Shemilt, A. Synnot, T. Turner, J. Elliott, T. Agoritsas, J. Hilton, C. Perron, E. Akl, R. Hodder, C. Pestridge, L. Albrecht, T. Horsley, J. Platt, R. Armstrong, P.H. Nguyen, R. Plovnick, A. Arno, N. Ivers, G. Quinn, A. Au, R. Johnston, G. Rada, M. Bagg, A. Jones, P. Ravaud, C. Boden, L. Kahale, B. Richter, I. Boisvert, H. Keshavarz, R. Ryan, L. Brandt, S.A. Kolakowsky-Hayner, D. Salama, A. Brazinova, S.K. Nagraj, G. Salanti, R. Buchbinder, T. Lasserson, L. Santaguida, C. Champion, R. Lawrence, N. Santesso, J. Chandler, Z. Les, H.J. SchÃnemann, A. Charidimou, S. Leucht, I. Shemilt, R. Chou, N. Low, D. Sherifali, R. Churchill, A. Maas, R. Siemieniuk, M.C. Cnossen, H. MacLehose, M. Simmonds, M.-J. Cossi, M. Macleod, N. Skoetz, M. Counotte, K. Soares-Weiser, S. Craigie, R. Marshall, V. Srikanth, P. Dahm, N. Martin, K. Sullivan, A. Danilkewich, L.M. GarcÃa, K. Danko, M. Taylor, E. Donoghue, L.J. Maxwell, K. Thayer, C. Dressler, J. McAuley, C. Egan, R. Tritton, J. McKenzie, G. Tsafnat, S.A. Elliott, J. Meerpohl, P. Tugwell, I. Etxeandia, B. Merner, R. Featherstone, S. Mondello, R. Foxlee, R. Morley, G. van Valkenhoef, P. Garner, A. Turgeon, P. Vandvik, M. Gerrity, Z. Munn, M. Murano, S.A. Wallace, S. Green, K. Newman, C. Watts, J. Grimshaw, R. Nieuwlaat, L. Weeks, K. Gurusamy, A. Nikolakopoulou, A. Weigl, N. Haddaway, G. Wells, L. Hartling, A. O'Connor, W. Wiercioch, J. Hayden, M. Page, L. Wolfenden, M. Helfand, M. Pahwa, J.J. Yepes NuÃez, J. Higgins, J.P. Pardo, J. Yost, S. Hill, and L. Pearson. Living systematic reviews: 2. combining human and machine effort. *Journal of Clinical Epidemiology*, 91:31–37, November 2017. ISSN 0895-4356.

Jacob Walker and Philip Taylor. Using ebird data to model population change of migratory bird species. *Avian Conservation and Ecology*, 12(1), June 2017. ISSN 1712-6568.

Michael J Way, Ashok N Srivastava, Jeffrey D Scargle, and Kamal M Ali. *Advances in machine learning and data mining for astronomy*. Chapman and Hall/CRC, 2012.

C.O. Webb, J.W.F. Slik, and T. Triono. Biodiversity inventory and informatics in southeast asia. *Biodiversity and Conservation*, 19(4):955–972, April 2010. ISSN 0960-3115.

Darryl E. Wright, Chris J. Lintott, Stephen J. Smartt, Ken W. Smith, Lucy Fortson, Laura Trouille, Campbell R. Allen, Melanie Beck, Mark C. Bouslog, Amy Boyer, K. C. Chambers, Heather Flewelling, Will Granger, Eugene A. Magnier, Adam McMaster, Grant R. M. Miller, James E. O'Donnell, Brooke Simmons, Helen Spiers, John L. Tonry, Marten Veldthuis, Richard J. Wainscoat, Chris Waters, Mark Willman, Zach Wolfenbarger, and Dave R. Young. A transient search using combined human and machine classifications. *MONTHLY NOTICES OF THE ROYAL ASTRONOMICAL SOCIETY*, 472(2):1315–1323, December 2017. ISSN 0035-8711.

A.16 Volunteer Computing

Articles concerning Volunteer Computing initiatives:

Vickie Curtis. Patterns of participation and motivation in folding home: The contribution of hardware enthusiasts and overclockers. *Citizen Science: Theory and Practice*, 3(1), 2018.

A. Hornig, L. Bausch, and U. Beckert. Space science and society - motivating the masses to donate idle computing time at home for numerical aerospace research in distributed computing and citizen science. volume 13, pages 10944–10949, 2012.

Laure Kloetzer, Julien Da Costa, and Daniel K Schneider. Not so passive: Engagement and learning in volunteer computing projects. *Human Computation*, 3(1):25–68, 2016.

Joseph L Myers, Cory Lehan, Pamela Gay, Matthew Richardson, and CosmoQuest Team. Cosmoquest transient tracker: Opensource photometry & astrometry software. In *American Astronomical Society Meeting Abstracts*, volume 231, 2018.

O. Nov, D. Anderson, and O. Arazy. Volunteer computing: A model of the factors determining contribution to community-based scientific research. pages 741–750, 2010.

Victor I Tishchenko. The behavioral patterns of volunteer computing communities.

David Toth, Russell Mayer, and Wendy Nichols. Increasing participation in volunteer computing. In *Parallel and Distributed Processing Workshops and Phd Forum (IPDPSW), 2011 IEEE International Symposium on*, pages 1878–1882. IEEE, 2011.

Appendix B

Round Two Excluded Publications

B.1 Applications of VCS

Cathal Doyle, Yevgeniya Li, Markus Luczak-Roesch, Dayle Anderson, Brigitte Glasson, Matthew Boucher, Carol Brieseman, Dianne Christenson, and Melissa Coton. What is online citizen science anyway? an educational perspective. *arXiv preprint arXiv:1805.00441*, 2018.

B.2 Cannot Ensure Online

Anne Bowser, Katie Shilton, Jenny Preece, and Elizabeth Warrick. Accounting for privacy in citizen science: Ethical research in a context of openness. In *Proceedings of the 20th ACM Conference on Computer-Supported Cooperative Work and Social Computing (CSCW 2017)*, pages 2124–2136, 2017.

Anne Bowser and Andrea Wiggins. Privacy in participatory research: Advancing policy to support human computation. *Human Computation*, 2(1):19–44, 2015.

Barbara Kieslinger, Teresa Schäfer, Florian Heigl, Daniel Dörler, Anett Richter, and Aletta Bonn. The challenge of evaluation: An open framework for evaluating citizen science activities. 2017.

Peter Mooney and Lorraine Morgan. How much do we know about the contributors to volunteered geographic information and citizen science projects? *ISPRS Annals of Photogrammetry, Remote Sensing & Spatial Information Sciences*, 2, 2015.

Sarah Elizabeth West and Rachel Mary Pateman. Recruiting and retaining participants in citizen science: What can be learned from the volunteering literature? *Citizen Science: Theory and Practice*, 2016.

B.3 Inaccessible

Devreni-Koutsouki Anna, Dobreva Milena, and Jennings Edel. Towards user engagement models for citizen science: Initiatives in the digital cultural heritage domain. In *Cultural Heritage Communities*, pages 78–95. Routledge, 2017.

Francesco Cappa, Jeffrey Laut, Luca Giustiniano, Oded Nov, and Maurizio Porfiri. Interactions with researchers in online citizen science projects. In *XXVI Riunione Scientifica Annuale Associazione Italiana di Ingegneria Gestionale (RSA AiIG 2015)*, pages 1–14, 2015.

Charlene Jennett and Anna L Cox. Digital citizen science and the motivations of volunteers. *The Wiley Handbook of Human Computer Interaction*, 2:831–841, 2018.

CHARLENE Jennett, LAURE Kloetzer, Anna L Cox, Daniel Schneider, EMILY Collins, MATTIA Fritz, MICHAEL J Bland, CINDY Regalado, Ian Marcus, HANNAH Stockwell, et al. Creativity in citizen cyberscience. *Human Computation*, 3(1):181–204, 2017.

Edel Jennings, Milena Dobreva, and Anna Devreni-Koutsouki. Towards user engagement models for citizen science: initiatives in the digital cultural heritage domain. In *Cultural Heritage Communities*, pages 98–115. Routledge, 2017.

Jeffrey Laut, Francesco Cappa, Oded Nov, and Maurizio Porfiri. A model for citizen scientist contribution in an image tagging task. In *ASME 2016 Dynamic Systems and Control Conference*, volume 1, pages V001T14A001–V001T14A001. American Society of Mechanical Engineers, 2016. ISBN 978-0-7918-5069-5. 9th ASME Annual Dynamic Systems and Control Conference, Minneapolis, MN, OCT 12-14, 2016.

Bruce V Lewenstein. Two meanings of citizen science caren b. cooper and. *The Rightful Place of Science*, page 51.

Julia K Parrish, Hillary Burgess, Jake F Weltzin, Lucy Fortson, Andrea Wiggins, and Brooke Simmons. Exposing the science in citizen science: Fitness to purpose and intentional design. *Integrative and comparative biology*, 2018.

J.F. Sherson. Studying human problem solving using citizen science games. pages 320–325, 2017.

B.4 Incomplete

M. Aristeidou, E. Scanlon, and M. Sharples. A design-based study of citizen inquiry for geology. volume 1093, pages 7–13, 2013.

Shanique Brown, PL Gay, and CS Daus. Motivation of citizen scientists participating in moon zoo. In *Bulletin of the American Astronomical Society*, volume 43, 2011.

Lucy Fortson and Stuart Lynn. Talking in the zooniverse: a collaborative tool for citizen scientists. In WW Smari, GC Fox, and M Nygard, editors, *Collaboration Technologies and Systems (CTS), 2014 International Conference on*, pages 1–2. IEEE, 2014. ISBN 978-1-4799-5158-1. International Conference on Collaboration Technologies and Systems (CTS), Minneapolis, MN, MAY 19-23, 2014.

Jennifer Hammock, Derek L Hansen, Anne Bowser, Jennifer Preece, Carol Boston, and Yurong He. Towards citizen inquiry: From class-based environmental projects to citizen science. In *Citizen Inquiry*, pages 143–167. Routledge, 2017.

Katie DeVries Hassman, Gabriel Mugar, Corey Jackson, et al. Learning at the seafloor, looking at the sky: The relationship between individual tasks and collaborative engagement in two citizen science projects. volume 2, pages 265–266. International Society of the Learning Sciences, 2013.

Gitte Kragh. The motivations of volunteers in citizen science. *environmental SCIENTIST*, 25(2):32–35, 2016.

Roman Lukyanenko and Jeffrey Parsons. Beyond task-based crowdsourcing database research. In *AAAI Conference on Human Computation & Crowdsourcing (AAAI HCOMP)*, pages 1–2, 2015.

ROMAN LUKYANENKO, JEFFREY PARSONS, and YOLANDA F WIERSMA. The impact of task design on accuracy, completeness and discovery in surveillance-based crowdsourcing. In *Proceedings of the 4th International Conference on Collective Intelligence*, 2016.

Vineet Pandey, Scott Klemmer, Amnon Amir, Justine Debelius, Embriette R Hyde, Tomasz Kosciolek, and Rob Knight. Integrating citizen science with online learning to ask better questions. *arXiv preprint arXiv:1609.05763*, 2016.

J. Reed, W. Rodriguez, and A. Rickhoff. A framework for defining and describing key design features of virtual citizen science projects. pages 623–625, 2012.

Jason T Reed, Ryan Cook, M Jordan Raddick, Karen Carney, and Chris Lintott. Participating in online citizen science: Motivations as the basis for user types and trajectories. In *Handbook of Human Computation*, pages 695–702. Springer, 2013.

Robert Simpson. Citizen science: Turning everyone into a scientist. *TRANSACTIONS OF THE LEICESTER LITERARY & PHILOSOPHICAL SOCIETY.*, page 27, 2011.

James Sprinks. Human factors in citizen science.

Claudia Viggiano. Community-specific language in online citizen science forums: a corpus-driven diachronic study.

Andrea Wiggins. Crowdsourcing science: organizing virtual participation in knowledge production. In *Proceedings of the 16th ACM international conference on Supporting group work*, pages 337–338. ACM, 2010.

B.5 Little Contribution

Sonia Fizek. All work no play: are games becoming the factories of the future? *www. firstpersonscholar. com*, 2016.

Christothea Herodotou, Mike Sharples, and Eileen Scanlon. *Citizen Inquiry: Synthesising Science and Inquiry Learning*. Routledge, 2017.

Trent A Mankowski, Stephanie J Slater, and Timothy F Slater. An interpretive study of meanings citizen scientists make when participating in galaxy zoo. *Contemporary Issues in Education Research*, 4(4):25–42, 2011.

Emma Weitkamp. From planning to motivations: citizen science comes to life. *Journal of Science Communication*, 15(3), 2016.

Poonam Yadav and John Darlington. Design guidelines for the user-centred collaborative citizen science platforms. *arXiv preprint arXiv:1605.00910*, 2016.

B.6 Not Peer Reviewed - Editorial/Opinion

Anne Bowser, Andrea Wiggins, Lea Shanley, Jennifer Preece, and Sandra Henderson. Sharing data while protecting privacy in citizen science. *interactions*, 21(1):70–73, 2014.

R. Khelifa. Wildlife monitoring: Lure gamers into citizen science. *Nature*, 537 (7621):488, 2016.

Roman Lukyanenko, Jeffrey Parsons, and Yolanda F Wiersma. Emerging problems of data quality in citizen science. *Conservation Biology*, 30(3): 447–449, 2016.

D.J. Tippins and L.J. Jensen. Citizen science in digital worlds: The seduction of a temporary escape or a lifelong pursuit? *Cultural Studies of Science Education*, 7(4):851–856, 2012.

Kathleen Toerpe. The rise of citizen science. *The Futurist*, 47(4):25, 2013.

B.7 Not Peer Reviewed - Poster/Presentation

Cassie BOWMAN. An overview of the state of online citizen science and crowdsourcing. In *2010 GSA Denver Annual Meeting*, 2010.

G Bracey, A Glushko, MN Bakerman, P Gay, and S Buxner. Cosmoquest: Exploring the needs of current & future citizen scientists. In *AGU Fall Meeting Abstracts*, 2016.

S DE FELICI, F Sorge, V Sbordoni, and D Cesaroni. Scientists by chance: what tell us data from unaware citizen scientists? In *First International ECSA Conference 2016-Citizen Science–Innovation in Open Science, Society and Policy*. DE, 2016.

Joseph Moore, PL Gay, K Hogan, C Lintott, C Impey, and C Watson. Facebooking citizen science with the zooniverse. In *Bulletin of the American Astronomical Society*, volume 43, 2011.

Maya Nona Bakerman, Sanlyn Buxner, Georgia Bracey, and Nicole Gugliucci. Assessing motivations and use of online citizen science astronomy projects. In *American Astronomical Society Meeting Abstracts*, volume 231, 2018.

Aaron Price. The use of the nature of scientific knowledge scale as a entrance assessment in a large, online citizen science project. In *Bulletin of the American Astronomical Society*, volume 42, page 509, 2010.

Jordan Raddick, GL Bracey, and PL Gay. Motivations of citizen scientists participating in galaxy zoo: A more detailed look. In *Bulletin of the American Astronomical Society*, volume 42, page 509, 2010.

J. Radford, A. Pilny, K. Ognyanova, L. Horgan, S. Wojcik, and D. Lazer. Gaming for science: A demo of online experiments on volunteerscience.com. volume 26-February-2016, pages 86–89, 2016.

Cia Romano, Pamela Gay, Ryan Owens, and Grigori Burlea. Why citizen science without usability testing will underperform. In *AGU Fall Meeting Abstracts*, 2017.

SJ Slater, T Mankowski, TF Slater, et al. Dreamers, poets, citizens, and scientists: Motivations for engaging in galaxyzoo citizen science. In *AGU Fall Meeting Abstracts*, 2010.

Jian Tang and Nathan R Prestopnik. Effects of framing on user contribution: Story, gameplay and science. *AMCIS 2017 - America's Conference on Information Systems: A Tradition of Innovation*, 2017-August, 2017.

B.8 Not Peer Reviewed - Technical Report

Rick Bonney, Heidi Ballard, Rebecca Jordan, Ellen McCallie, Tina Phillips, Jennifer Shirk, and Candie C Wilderman. Public participation in scientific research: Defining the field and assessing its potential for informal science education. a caise inquiry group report. *Online Submission*, 2009.

Anne Bowser and Lea Shanley. New visions in citizen science. *Case Study Ser*, 3:1–53, 2013.

Karl Clarke. Defining one's role in citizen science: an exploration of the roles, perceptions and outcomes of participation in citizen science activities. 2012.

Monica Duke and Emma Tonkin. Patients participate! literature review: Usability and human factors in citizen science projects, and trust and credibility on the web. 2012.

H Geoghegan, A Dyke, R Pateman, S West, and G Everett. Understanding motivations for citizen science. *Final report on behalf of UKEOF, University of Reading, Stockholm Environment Institute (University of York) and University of the West of England*, 2016.

Muki Haklay. Citizen science and policy: a european perspective. *The Wodrow Wilson Center, Commons Lab*, 2015.

E. Hand. Citizen science: People power. *Nature*, 466(7307):685–687, 2010.

Eleftheria Vasileiadou. Crowd science: it is not just a matter of time (or funding). *Journal of the Association for Information Science and Technology*, 66 (7):1514–1517, 2015.

SA Wigboldus. Considering the potential of citizens' science. 2016.

B.9 Not Peer Reviewed - Thesis

Mark D Cottman-Fields. *Virtual birding: extending birdwatching to review acoustic recordings*. PhD thesis, Queensland University of Technology, 2017.

Dehn Giuliana. Designing an interface for citizen science platforms ensuring a good user experience. 2017.

Ruth Kermish-Allen. Designing for online collaborations and local environmental action in citizen science: A multiple case study. 2016.

Erica Krimmel and Lili Luo. Keeping the citizens scientists: Participant motivations in citizen science. 2013.

RJ Murphy. Facilitating citizen science through gamification. *Department of Design Science and Information Systems, Florida International University*, 2015.

M Prats Lopez. Managing citizen science in the humanities: The challenge of ensuring quality. 2017.

Alison Seeberger. There's no such thing as free labor: Evaluating citizen science volunteer motivations. 2014.

Jonathan A Yee. Characterizing crowd participation and productivity of foldit through web scraping. Technical report, Naval Postgraduate School Monterey United States, 2016.

B.10 Not Peer Reviewed - Workshop

G Bracey, K Costello, P Gay, and E Reilly. Citizen science: Mapping the moon and mercury. In JB Jensen, JG Manning, MG Gibbs, and D Daon, editors, *Connecting People to Science: A National Conference on Science Education and Public Outreach*, volume 457 of *Astronomical Society of the Pacific Conference Series*, page 93. Space Telescope Sci Inst; NASA Lunar Sci Inst; NASAs Explorat Program; Infrared Process & Anal Ctr; NASAs Herschel Sci Ctr; Spitzer Sci Ctr; Stratospher Observ Infrared Astron; NASAs Chandra XRay Observ; Univ Chicago Press; Natl Radio Astron Observ; Ball Aerospace; Capitol Coll; Sky Skan; Univ Wyoming CAPER Team; Amer Astron Soc; AAS Educ Off; Solar Dynam Observ; Seiler Instrument; Explore Sci; Pratt St Ale House; MWT Assoc Inc; Amer Elements; Charlesbridge; Celestron, 2012. ISBN 978-1-58381-796-4. 123rd Annual Meeting of the Astronomical-Society-of-the-Pacific, Amer Geophys Union, Baltimore, MD, JUL 30-AUG 03, 2011.

Jana B Jarecki. Citizen science, gamification, and virtual reality for cognitive research.

C. Jennett and A.L. Cox. Eight guidelines for designing virtual citizen science projects. volume WS-14-20, pages 16–17, 2014.

Jessica L Oliver, Mark Cottman-Fields, Margot Brereton, and Paul Roe. Co-designing technologies for citizen scientist engagement & saving species. In *Proceedings of the 29th Australian Conference on Computer-Human Interaction*, pages 651–653. ACM, 2017.

Andrea Wiggins and Kevin Crowston. Distributed scientific collaboration: Research opportunities in citizen science. In *Proceedings of the CSCW 2010 workshop on Changing Dynamics of Scientific Collaboration*, 2010.

B.11 Not VCS - Offline

Abdulmonem Alabri and Jane Hunter. Enhancing the quality and trust of citizen science data. In *e-Science (e-Science), 2010 IEEE Sixth International Conference on*, pages 81–88. IEEE, 2010.

Marcio Antelio, Maria Gilda P Esteves, Daniel Schneider, and Jano Moreira de Souza. Qualitocracy: A data quality collaborative framework applied to citizen science. In *Systems, Man, and Cybernetics (SMC), 2012 IEEE International Conference on*, pages 931–936. IEEE, 2012.

Willemijn M Appels, Lori Bradford, Kwok P Chun, Anna E Coles, and Graham Strickert. Diy meteorology: Use of citizen science to monitor snow dynamics in a data-sparse city, 2017.

Maria Aristeidou, Eileen Scanlon, and Mike Sharples. Design processes of a citizen inquiry community. *Citizen Inquiry: Synthesising Science and Inquiry Learning*, pages 210–229, 2017.

Anne Bowser, Derek Hansen, Yurong He, Carol Boston, Matthew Reid, Logan Gunnell, and Jennifer Preece. Using gamification to inspire new citizen science volunteers. In *Proceedings of the first international conference on gameful design, research, and applications*, pages 18–25. ACM, 2013.

C. Brooking and J. Hunter. Reputation-aware filtering services for citizen science data. pages 7–13, 2011.

Caren Cooper. *Citizen science: How ordinary people are changing the face of discovery*. Gerald Duckworth & Co, 2017.

Mark Cottman-Fields, Margot Brereton, and Paul Roe. Virtual birding: extending an environmental pastime into the virtual world for citizen science. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 2029–2032. ACM, 2013.

M. DolejÃ and D. Kera. Soylent diet self-experimentation: Design challenges in extreme citizen science projects. pages 2112–2123, 2017.

Elizabeth R Ellwood, Betty A Dunckel, Paul Flemons, Robert Guralnick, Gil Nelson, Greg Newman, Sarah Newman, Deborah Paul, Greg Riccardi, Nelson Rios, et al. Accelerating the digitization of biodiversity research specimens through online public participation. *BioScience*, 65(4):383–396, April 2015. ISSN 0006-3568.

Troy Frensley, Alycia Crall, Marc Stern, Rebecca Jordan, Steven Gray, Michelle Prysby, Greg Newman, Cindy Hmelo-Silver, David Mellor, and Joey Huang. Bridging the benefits of online and community supported citizen science: A case study on motivation and retention with conservation-oriented volunteers. *Citizen Science: Theory and Practice*, 2(1), 2017.

K.K. Fuccillo, T.M. Crimmins, C.E. de Rivera, and T.S. Elder. Assessing accuracy in citizen science-based plant phenology monitoring. *International Journal of Biometeorology*, 59(7):917–926, July 2015. ISSN 0020-7128.

M. Gharesifard and U. Wehn. To share or not to share: Drivers and barriers for sharing data via online amateur weather networks. *Journal of Hydrology*, 535: 181–190, April 2016a. ISSN 0022-1694.

M. Gharesifard and U. Wehn. *What drives citizens to engage in ICT-enabled citizen science? Case study of online amateur weather networks*. Advances in Knowledge Acquisition Transfer and Management (AKATM) Book Series. 2016b. ISBN 978-1-5225-0963-9; 978-1-5225-0962-2.

M. Gharesifard, U. Wehn, and P. van der Zaag. Towards benchmarking citizen observatories: Features and functioning of online amateur weather networks. *Journal of Environmental Management*, 193:381–393, may 15 2017. ISSN 0301-4797.

Barbara Haltofova. Leveraging collective intelligence of online users for productive outcomes. In S Moffett and B Galbraith, editors, *PROCEEDINGS OF THE 17TH EUROPEAN CONFERENCE ON KNOWLEDGE MANAGEMENT*, Proceedings of the European Conference on Knowledge Management, pages 1031–1037, 2016. ISBN 978-1-911218-03-6. 17th European Conference on Knowledge Management (ECKM), Ulster Univ, NORTH IRELAND, SEP 01-02, 2016.

Y. He, J. Preece, C. Boston, A. Bowser, D. Hansen, and J. Hammock. The effects of individualized feedback on college students' contributions to citizen science. pages 165–168, 2014.

Y. He, M. Weber, J. Preece, S. McKeon, J. Hammock, and A. Wiggins. A journey of citizen science data in an online environment. volume 26-February-2016, pages 289–292, 2016.

Pekka Helle, Katja Ikonen, and Anu Kantola. Wildlife monitoring in finland: online information for game administration, hunters, and the wider public. *CANADIAN JOURNAL OF FOREST RESEARCH*, 46(12):1491–1496, December 2016. ISSN 0045-5067. 17th Conference of the International-Boreal-Forest-Research- Association (IBFRA), Rovaniemi, FINLAND, MAY 24-29, 2015.

Wesley M Hochachka, Alison Johnston, Steve Kelling, Travis Moore, Weng-Keen Wong, Jun Yu, et al. Can observation skills of citizen scientists be estimated using species accumulation curves? 2015.

J. Huang, C.E. Hmelo-Silver, R. Jordan, S. Gray, T. Frensley, G. Newman, and M.J. Stern. Scientific discourse of citizen scientists: Models as a boundary object for collaborative problem solving. *Computers in Human Behavior*, 2018.

Jane Hunter, Abdulmonem Alabri, and Catharine van Ingen. Assessing the quality and trustworthiness of citizen science data. *Concurrency and Computation: Practice and Experience*, 25(4):454–466, 2013. ISSN 1532-0626.

Ruth Kermish-Allen. Design principles of online learning communities in citizen science. *Maine Policy Review*, 26(2): 80–85, 2017.

Sunyoung Kim. A flexible tool for participating, authoring, and managing citizen science campaigns on-the-go. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, pages 556–559. ACM, 2012.

Sunyoung Kim, Jennifer Mankoff, and Eric Paulos. Sensr: evaluating a flexible framework for authoring mobile data-collection tools for citizen science. In *Proceedings of the 2013 conference on Computer supported cooperative work*, pages 1453–1462. ACM, 2013.

Sunyoung Kim and Eric Paulos. A subscription-based authoring tool for mobile citizen science campaigns. In *CHI'12 Extended Abstracts on Human Factors in Computing Systems*, pages 2135–2140. ACM, 2012.

Anne M Land-Zandstra, Jeroen LA Devilee, Frans Snik, Franka Buurmeijer, and Jos M van den Broek. Citizen science on a smartphone: Participants' motivations and learning. *Public Understanding of Science*, 25(1):45–60, 2016.

Andrea Liberatore, Erin Bowkett, Catriona J MacLeod, Eric Spurr, and Nancy Longnecker. Social media as a platform for a citizen science community of practice. *Citizen Science: Theory and Practice*, 3(1), 2018.

Yong Liu, Denzil Ferreira, Jorge Goncalves, Simo Hosio, Pratyush Pandab, and Vassilis Kostakos. Donating context data to science: The effects of social signals and perceptions on action-taking. *Interacting with Computers*, 29(2):132–146, 2017.

Roman Lukyanenko and Jeffrey Parsons. Data quality as an outcome of conceptual modeling choices. In *16th the International Conference of Information Quality (ICIQ2011)*, 2011a.

Roman Lukyanenko and Jeffrey Parsons. Rethinking data quality as an outcome of conceptual modeling choices. In *16th International Conference on Information Quality*, pages 1–16, 2011b.

Roman Lukyanenko and Jeffrey Parsons. Extending participatory design principles to structured user-generated content. In *Scandinavian Conference on Information Systems*, pages 237–252. Springer, 2015a.

Roman Lukyanenko and Jeffrey Parsons. Principles for modeling user-generated content. In P Johannesson, ML Lee, SW Liddle, AL Opdahl, and OP Lopez, editors, *International Conference on Conceptual Modeling*, volume 9381 of *Lecture Notes in Computer Science*, pages 432–440. Springer, 2015b. ISBN 978-3-319-25264-3; 978-3-319-25263-6. 34th International Conference on Conceptual Modeling (ER), Stockholm, SWEDEN, OCT 19-22, 2015.

Roman Lukyanenko, Jeffrey Parsons, and Yolanda Wiersma. Citizen science 2.0: Data management principles to harness the power of the crowd. In H Jain, AP Sinha, and P Vitharana, editors, *International Conference on Design Science Research in Information Systems*, volume 6629 LNCS of *Lecture Notes in Computer Science*, pages 465–473. Springer, 2011. ISBN 978-3-642-20632-0. 6th International Conference on Design Science Research in Information Systems and Technology (DESRIST), Milwaukee, WI, MAY 05-06, 2011.

Roman Lukyanenko, Jeffrey Parsons, Yolanda F Wiersma, Renee Sieber, and Mahed Maddah. Participatory design for user-generated content: Understanding the challenges and moving forward. *Scandinavian J. Inf. Systems*, 28(1):2, 2016a.

ROMAN Lukyanenko, Y Wiersma, and JEFFREY Parsons. Is crowdsourced attribute data useful in citizen science? a study of experts and machines. *Collective Intelligence*, 2016b.

Rikke Magnussen. Involving lay people in research and professional development through gaming: A systematic mapping review. In *European Conference on Games Based Learning*, pages 394–401. Academic Conferences International Limited, 2017.

Victoria Y Martin. Citizen science as a means for increasing public engagement in science: presumption or possibility? *Science Communication*, 39(2):142–168, 2017.

A.B. Mattos, R.G. Herrmann, K.K. Shigeno, and R.S. Feris. A mission-oriented citizen science platform for efficient flower classification based on combination of feature descriptors. volume 1222, pages 45–52, 2014.

G. McCrory, C. Veeckman, and L. Claeys. Citizen science is in the air – engagement mechanisms from technology-mediated citizen science projects addressing air pollution. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 10673 LNCS:28–38, 2017.

F. Montargil and V. Santos. Communication with citizens in the first eu citizen observatories experiences. volume Part F129463, pages 96–103, 2017.

Mikołaj Morzy. Ict services for open and citizen science. *World Wide Web*, 18(4): 1147–1161, 2015.

G. Newman, D. Zimmerman, A. Crall, M. Laituri, J. Graham, and L. Stapel. User-friendly web mapping: Lessons from a citizen science website. *International Journal of Geographical Information Science*, 24(12):1851–1869, 2010. ISSN 1365-8816.

Greg Newman, Jim Graham, Alycia Crall, and Melinda Laituri. The art and science of multi-scale citizen science support. *Ecological Informatics*, 6(3-4): 217–227, 2011.

Oded Nov, Ofer Arazy, Kelly Lotts, and Thomas Naberhaus. Motivation-targeted personalized ui design: a novel approach to enhancing citizen science participation. In *ECSCW 2013: Proceedings of the 13th European Conference on Computer Supported Cooperative Work, 21-25 September 2013, Paphos, Cyprus*, pages 287–297. Springer, 2013.

Barbara Prainsack. Understanding participation: The `citizen science' of genetics. In B Prainsack, S Schicktanz, and G WernerFelmayer, editors, *GENETICS AS SOCIAL PRACTICE: TRANSDISCIPLINARY VIEWS ON SCIENCE AND CULTURE*, Theory Technology and Society, pages 147–164. 2014. ISBN 978-1-315-58430-0; 978-1-4094-5548-6.

Jennifer Preece. Citizen science: New research challenges for human–computer interaction. *International Journal of Human-Computer Interaction*, 32(8): 585–612, 2016.

Jennifer Preece, Carol Boston, Mary Lou Maher, Kazjon Grace, and Tom Yeh. From crowdsourcing technology design to participatory design and back again! In C Bernadas and D Minchella, editors, *PROCEEDINGS OF THE 3RD EUROPEAN CONFERENCE ON SOCIAL MEDIA*, Proceedings of the European Conference on Social Media, pages 315–323, 2016. ISBN 978-1-911218-01-2. 3rd European Conference on Social Media (ECSM), Ecole Management Normandie, Caen, FRANCE, JUL 12-13, 2016.

Chris Preist, Elaine Massung, and David Coyle. Competing or aiming to be average?: normification as a means of engaging digital volunteers. In *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing*, pages 1222–1233. ACM, 2014.

C. Robson, M.A. Hearst, C. Kau, and J. Pierce. Comparing the use of social networking and traditional media channels for promoting citizen science. pages 1463–1468, 2013.

N. Sharma. Species identification in citizen science: Effects of interface design and image difficulty on user performance and workload. volume 07-12-May-2016, pages 128–133, 2016.

Rejane Spitz, Clorisval Pereira Jr, Francisco Queiroz, Leonardo C Leite, Peter Dam, Marcelo P Ferranti, Renan Kogut, and Wesley Oliveira. Gamification, citizen science and civic engagement: in search of the common good.

Jared Starr, Charles M Schweik, Nathan Bush, Lena Fletcher, Jack Finn, Jennifer Fish, and Charles T Bargeron. Lights, camera... citizen science: assessing the effectiveness of smartphone-based video training in invasive plant identification. *PloS one*, 9(11):e111433, 2014.

Ulrike Sturm, M Gold, Soledad Luna, S Schade, L Ceccaroni, CCM Kyba, B Claramunt, M Haklay, D Kasperowski, A Albert, et al. Defining principles for mobile apps and platforms development in citizen science. *Research Ideas and Outcomes*, 2018.

Sarah-Kristin Thiel, Titiana Petra Ertiö, and Matthias Baldauf. Why so serious? the role of gamification on motivation and engagement in e-participation. *INTERACTION DESIGN AND ARCHITECTURES*, (35):158–181, 2017.

Brian Traynor, Tracy Lee, and Danah Duke. Case study: Building ux design into citizen science applications. In *International Conference of Design, User Experience, and Usability*, pages 740–752. Springer, 2017.

R. van der Wal, N. Sharma, C. Mellish, A. Robinson, and A. Siddharthan. The role of automated feedback in training and retaining biological recorders for citizen science. *Conservation Biology*, 30 (3):550–561, June 2016. ISSN 0888-8892.

Andrea Wiggins and Yurong He. Community-based data validation practices in citizen science. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*, volume 27, pages 1548–1559. ACM, 2016. ISBN 978-1-4503-3592-8. 19th ACM Conference on Computer-Supported Cooperative Work and Social Computing (CSCW), San Francisco, CA, FEB 27-MAR 02, 2016.

Poonam Yadav and John Darlington. Conceptual frameworks for building online citizen science projects. *arXiv preprint arXiv:1704.05084*, 2017.

J. Yu, S. Kelling, J. Gerbracht, and W.-K. Wong. Automated data verification in a large-scale citizen science project: A case study. 2012.

B.12 Not VCS - Other Crowdsourcing

J.-W. Ahn, J. Hammock, C. Parr, J. Preece, B. Shneiderman, K. Schulz, D. Hansen, D. Rotman, and Y. He. Visually exploring social participation in encyclopedia of life. pages 149–156, 2013.

Maria Aristeidou, Eileen Scanlon, and Mike Sharples. Profiles of engagement in online communities of citizen science participation. *Computers in Human Behavior*, 74:246–256, September 2017. ISSN 0747-5632.

Martina Balestra, Coye Cheshire, Ofer Arazy, and Oded Nov. Investigating the motivational paths of peer production newcomers. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pages 6381–6385. ACM, 2017a.

Martina Balestra, Lior Zalmanson, Coye Cheshire, Ofer Arazy, and Oded Nov. It was fun, but did it last? the dynamic interplay between fun motives and contributors' activity in peer-production. *Human-Computer Interaction*, 1:1, 2017b.

Matthias Budde, Andrea Schankin, Julien Hoffmann, Marcel Danz, Till Riedel, and Michael Beigl. Participatory sensing or participatory nonsense?: Mitigating the effect of human error on data quality in citizen science. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 1(3):39, 2017.

M. Dobreva. Collective knowledge and creativity: The future of citizen science in the humanities. *Advances in Intelligent Systems and Computing*, 416: 565–573, 2016. ISSN 2194-5357. 9th International Conference on Knowledge, Information and Creativity Support Systems (KICSS), Limassol, CYPRUS, NOV 06-08, 2014.

Libby Ellwood, Henry Bart, Michael Doosey, Dean Jue, Gil Nelson, Nelson Rios, and Austin Mast. Mapping life: Quality assessment of novice and computer automated vs. expert georeferences. In *TDWG 2015 ANNUAL CONFERENCE*, 2015.

Mordechai E Haklay. Why is participation inequality important? Ubiquity Press, 2016.

Kyungsik Han, Eric A Graham, Dylan Vassallo, and Deborah Estrin. Enhancing motivation in a mobile participatory sensing project through gaming. In *Privacy, Security, Risk and Trust (PASSAT) and 2011 IEEE Third Inernational Conference on Social Computing (SocialCom), 2011 IEEE Third International Conference on*, pages 1443–1448. IEEE, 2011.

D-Lib Magazine. Collaborative construction of digital cultural heritage: a synthesis of research on online sociability determinants. *D-Lib Magazine*, 21 (11/12), 2015.

Marjolaine Matabos, Maia Hoeberechts, Carol Doya, Jacopo Aguzzi, Jessica Nephin, Thomas E Reimchen, Steve Leaver, Roswitha M Marx, Alexandra Branzan Albu, Ryan Fier, et al. Expert, crowd, students or algorithm: who holds the key to deep-sea imagery 'big data' processing? *Methods in Ecology and Evolution*, 8(8):996–1004, August 2017. ISSN 2041-210X.

N. Oliveira, E. Jun, and K. Reinecke. Citizen science opportunities in volunteer-based online experiments. In *PROCEEDINGS OF THE 2017 ACM SIGCHI CONFERENCE ON HUMAN FACTORS IN COMPUTING SYSTEMS (CHI'17)*, volume 2017-May, pages 6800–6812. Assoc Comp Machinery; ACM SIGCHI, 2017. ISBN 978-1-4503-4655-9. ACM SIGCHI Conference on Human Factors in Computing Systems (CHI), Denver, CO, MAY 06-11, 2017.

B.13 Use of VCS as Methodology for Experiment

G.E. Austen, M. Bindemann, R.A. Griffiths, and D.L. Roberts. Species identification by conservation practitioners using online images: Accuracy and agreement between experts. *PeerJ*, 2018 (1), jan 25 2018. ISSN 2167-8359.

B.14 Volunteer Computing

Lo Lee. Embracing diversity: The exploration of user motivations in citizen science astronomy projects. In *American Astronomical Society Meeting Abstracts# 232*, volume 232, 2018.

Appendix C

Recommendations for Project Review

Table 22: Summary of themes and recommendations extracted from the online communities literature.

Theme	**Summary**	**Effect**
Task Visibility	Tools for tracking/finding work	Positive
	View others have complied	Positive (increase chance to comply)
Requests	Ask people to perform personalised tasks	Positive (increase chance to do work)
	Make requests to specific people	Positive (increased chance to comply)
	Keep requests simple and short	Positive (increased rate of compliance)
	Requests should come from high-status people	Positive (increased rate of contribution)
	Requests should come from popular figures	Positive (increased rate of compliance)
Framing	Messages should stress benefits of contribution	Positive (larger effect on intrinsically motivated individuals)
	Fear campaigns effective and persuasive	Positive (increased rate of contribution)
	Fear campaigns result in participants' reevaluating campaigns	Neutral-Negative (may be less effective for some participants)
	Participants should feel they have something unique to offer	Positive (increased rate of contribution)
Goals	Goals should be specific and challenging	Positive (increased rate of contribution)
	Goals should be coupled with deadlines	Positive (increasing rate of contribution as deadline nears)
	Goals should be coupled with feedback	Positive (increased rate of contribution)
Social Interaction	Contributions should be combined with social interaction	Positive (increased rate of contribution)

(Continued)

Theme	Summary	Effect
Feedback	Provide participants with sincere, performance-based feedback	Positive (enhanced motivation, particularly where positive, but only if sincere)
	Provide systematic, quantitative feedback	Positive (increased likelihood of verbal feedback)
	Comparative feedback functional only if high performance desirable and obtainable	Neutral/Positive (increased rate of contribution if high performance viewed as possible)
	Performance feedback leads to a game-like atmosphere	Neutral/Negative (Undesirable consequences for some communities)
Rewards	Introduce any form of reward	Positive/Negative (motivates contributions - may not be the case for VCS - may lead to low effort behaviour and cheating)
	Introduce task-contingent feedback	Positive/Negative (effective for larger tasks, otherwise motivates increased rates of contribution but not effort)
	Rewards should be performance-contingent	Positive (Task-contingent leads to low-effort contributions)
	Performance-contingent rewards must be designed carefully	Positive (Further discourages low-effort contributions)
	Introduce status/privilege rewards	Positive (Low chance of most active participants engaging in low-effort behaviours)
	Ensure eligibility and reward policies are unpredictable and nontransparent	Positive (Low chance of low-effort behaviours)
	Avoid task-contingent feedback for intrinsically motivating tasks	Negative (Reduction in intrinsic motivations, particularly for monetary rewards)
	Introduce large tangible rewards for intrinsically motivating tasks	Positive (Small rewards have a negative impact on contribution rates)
Commitment	Encourage commitment to community groups	Positive (Increased willingness to contribute)
	Keep group sizes small	Positive (Greater effect than smaller groups)

Appendix D

Final List of Projects for Project Review

Project Name	Project URL	Community Features
Annotate	https://anno.tate.org.uk	Talk (Zooniverse Custom Platform)
Asteroid Mappers	https://cosmoquest.org/?application=vesta_mappers/	Discussion Board Forum
Asteroid Zoo	http://www.asteroidzoo.org	Talk (Zooniverse Custom Platform)
Bug Guide	http://bugguide.net/	Discussion Board Forum
Chicago Wildlife Watch	http://www.chicagowildlifewatch.org	Talk (Zooniverse Custom Platform)
Chimp and See	http://www.chimpandsee.org	Talk (Zooniverse Custom Platform)
Condor Watch	http://www.condorwatch.org	Talk (Zooniverse Custom Platform)
Cyclone Centre	http://www.cyclonecenter.org	Talk (Zooniverse Custom Platform)
Disk Detective	http://www.diskdetective.org	Talk (Zooniverse Custom Platform)
EteRNA	http://www.eternagame.org	Live Instant Messenger Chat, Discussion Board Forum, Wiki
EyeWire	http://eyewire.org/	Live Instant Messenger Chat, Discussion Board Forum, Wiki

(Continued)

Project Name	Project URL	Community Features
Floating Forests	http://www.floatingforests.org	Talk (Zooniverse Custom Platform)
FoldIt	http://fold.it/portal/	Discussion Board Forum, Wiki
Fossil Finder	http://www.zooniverse.org/projects/adrianevans/fossil-finder/	Talk (Zooniverse Custom Platform)
Galaxy Zoo	http://www.galaxyzoo.org	Talk (Zooniverse Custom Platform)
Galaxy Zoo Bar Lengths	http://www.zooniverse.org/projects/vrooje/galaxy-zoo-bar-lengths/	Talk (Zooniverse Custom Platform)
Herbaria@Home	http://herbariaunited.org/atHome/	Discussion Board Forum
Higgs Hunters	http://www.higgshunters.org	Talk (Zooniverse Custom Platform)
Instant Wild	http://www.edgeofexistence.org/instantwild/	Comment Listing
iSpotNature	http://www.ispotnature.org	Discussion Board Forum
Landscape Watch Hampshire	http://www.hampshire.landscapewatch.com/	Discussion Board Forum
Mars Mappers	https://cosmoquest.org/?application=mars_simply_craters	Discussion Board Forum
Mercury Mappers	https://cosmoquest.org/projects/mercury_mappers	Discussion Board Forum
Milky Way Project	http://www.milkywayproject.org/	Talk (Zooniverse Custom Platform)
Moon Mappers	https://cosmoquest.org/?application=simply_craters	Discussion Board Forum
Notes From Nature	http://www.notesfromnature.org	Talk (Zooniverse Custom Platform)
Old Weather	http://www.oldweather.org	Discussion Board Forum
Operation War Diary	http://www.operationwardiary.org	Talk (Zooniverse Custom Platform)
Orchid Observers	http://www.orchidobservers.org	Talk (Zooniverse Custom Platform)
Penguin Watch	http://www.penguinwatch.org	Talk (Zooniverse Custom Platform)

(Continued)

Project Name	Project URL	Community Features
Phylo	http://phylo.cs.mcgill.ca/	Discussion Board Forum
Planet Four	http://www.planetfour.org	Talk (Zooniverse Custom Platform)
Planet Four: Terrains	http://www.zooniverse.org/projects/mschwamb/planet-four-terrains/	Talk (Zooniverse Custom Platform)
Planet Hunters	http://www.planethunters.org	Talk (Zooniverse Custom Platform)
Plankton Portal	http://planktonportal.org/	Talk (Zooniverse Custom Platform)
Radio Galaxy Zoo	http://radio.galaxyzoo.org	Talk (Zooniverse Custom Platform)
Science Gossip	http://www.sciencegossip.org	Talk (Zooniverse Custom Platform)
Season Spotter Image Marking	http://www.zooniverse.org/projects/kosmala/season-spotter-image-marking	Talk (Zooniverse Custom Platform)
Season Spotter Questions	http://www.zooniverse.org/projects/kosmala/season-spotter-questions	Talk (Zooniverse Custom Platform)
Snapshot Serengeti	http://snapshotserengeti.org/	Talk (Zooniverse Custom Platform)
SpaceWarps	http://spacewarps.org/	Talk (Zooniverse Custom Platform)
Stardust@Home	http://stardustathome.ssl.berkeley.edu/	Discussion Board Forum
Sunspotter	http://www.sunspotter.org	Talk (Zooniverse Custom Platform)
Verb Corner	http://gameswithwords.org/VerbCorner	Discussion Board Forum
Whales As Individuals	http://www.zooniverse.org/projects/tedcheese/whales-as-individuals	Talk (Zooniverse Custom Platform)
Wildcam Gorongosa	http://www.wildcamgorongosa.org	Talk (Zooniverse Custom Platform)
Wildebeest Watch	http://www.zooniverse.org/projects/aliburchard/wildebeest-watch/	Talk (Zooniverse Custom Platform)
Worm Watch Lab	http://www.wormwatchlab.org	Talk (Zooniverse Custom Platform)

Appendix E

Interview Question Schedule

E.1 Introductory Briefing

Thank you for taking part in this study. I would like to remind you that you are free to withdraw from it at any time.

E.2 Interview Questions

E.2.1 Introduction

8. Please could you confirm again that you are happy to be (audio) recorded for the duration of this interview?

9. Could you begin by introducing yourself and your role in the team?

10. What is your interaction with other members of the team like?

E.2.2 Use and Introduction of Project (Game) Elements

11. To begin, can you go over the most important game elements that you use in your project, and how effective you think they have been?

12. When introducing new features, can you describe whether you base your development lifecycle on features you know work in other games, or on your own insights about how players use your own platform?

13. Has this changed over time? If so, please discuss how? (For example, have you become more aware about which features work well, generally?)

14. Are there any features or elements that you've considered introducing but subsequently haven't done so? Why?

E.2.3 Feedback on Features

15. After implementing a new feature, what is your process for receiving feedback on it from the community?

16. How do you review that feedback and then act upon it?

17. Do you specifically ask the community for feedback?

18. Do you get 'feedback' from data regarding engagement levels?

19. What is the relationship between data analytics on the platform and the way in which you re-engineer features, compared to when you speak to the community?

20. Do you value one source of data more than the other for this, for example?

E.2.4 Analysing the Impact of Features

21. When looking at effects of features on user engagement and motivation, do you study this longitudinally or simply at single time points?

22. How do you track exactly how specific changes result in specific effects?

23. What would you define community engagement to be?

24. Do you use yrou observations to gain insights into community health?

25. When things do not go to plan, what processes do you follow to pinpoint the source of problems?

26. What mechanisms do you have to subsequently re-engage the community?

E.2.5 Successful Citizen Science and Successful

27. What do think makes a CS project successful?

    (c) Social Success and features

(d) Community Success and features

(e) Task design and task success

28. What are your measures for success?

(f) Qualitative Measures

(g) Quantitative Measures

29. What do you think makes a game successful?

30. Do all the aspects you've described apply to creating a successful Citizen Science game?

31. If not, why not?

From Crowd to Community: A Survey of Online Community Features in Citizen Science Projects

Neal Reeves
University of Southampton
Southampton, UK
ntr1g09@soton.ac.uk

Ramine Tinati
University of Southampton
Southampton, UK
r.tinati@soton.ac.uk

Sergej Zerr
University of Southampton
Southampton, UK
s.zerr@soton.ac.uk

Max Van Kleek
University of Oxford
Oxford, UK
emax@cs.ox.ac.uk

Elena Simperl
University of Southampton
Southampton, UK
e.simperl@soton.ac.uk

ABSTRACT
Online citizen science projects have been increasingly used in a variety of disciplines and contexts to enable large-scale scientific research. The successes of such projects have encouraged the development of customisable platforms to enable anyone to run their own citizen science project. However, the process of designing and building a citizen science project remains complex, with projects requiring both human computation and social aspects to sustain user motivation and achieve project goals. In this paper, we conduct a systematic survey of 48 citizen science projects to identify common features and functionality. Supported by online community literature, we use structured walkthroughs to identify different mechanisms used to encourage volunteer contributions across four dimensions: task visibility, goals, feedback, and rewards. Our findings contribute to the ongoing discussion on citizen science design and the relationship between community and microtask design for achieving successful outcomes.

ACM Classification Keywords
H.5.m. Information Interfaces and Presentation (e.g. HCI): Miscellaneous

Author Keywords
Citizen science, online communities, survey

INTRODUCTION
Online citizen science (CS) refers to a Web-based approach to involve members of the general public in scientific research on a volunteer basis [6, 47, 62]. CS projects are typically initiated and overseen by a team of professional scientists, who define the goals of the projects, assign tasks to volunteers, and feed the crowd-generated data into established scientific workflows. This emerging form of participatory research has been applied to a vast range of scenarios, including education, civic activism, and conservation [63], alongside a growing number of disciplines, from astrophysics and biology to social sciences and cultural heritage [36, 59].

CSCW '17 February 25 - March 01, 2017, Portland, OR, USA

ACM ISBN 978-1-4503-4335-0/17/03.
DOI: http://dx.doi.org/10.1145/2998181.2998302

Citizen science draws on methods and theories from the fields of Human-Computer Interaction (HCI) and Computer Supported Cooperative Work (CSCW). As data-processing systems, they can be considered human-agent collectives [16] that use human cooperative work and crowd intelligence [32] to help professional scientists handle large amounts of raw data and advance their empirical work. As socio-technical systems, they foster an environment which enables loose-knit communities to form [40, 62], whose members communicate and collaborate using discussion forums [5, 59], chat rooms [58], or wikis [30].

The combination of human computation and sociality has shown to be effective in accomplishing scientific objectives, as well as in yielding unanticipated discoveries initiated by members of the community [59]. However, support for such communities varies significantly between different CS projects. Designing CS projects remains a complex process, requiring insight from a range of specialised areas and disciplines [3]. At the same time, the number of scientists involved can be small, lacking the prerequisite knowledge and expertise from areas such as HCI and CSCW, in order to design and run projects, and lack experience in successfully staging public engagement activities [59].

The motivation behind this paper is driven by the belief that design guidelines are crucial to simplifying this design process in citizen science projects. Whilst there has been focus on concentrating on factors such as participant activity levels as measure for successful CS initiatives [59], literature has begun to reveal the significance of community-specific project features, yet they still remains largely under-explored beyond a single project. Addressing this gap, this paper contains a study conducted investigating the features commonly used in 48 citizen science projects. We drew from an initial set of 136 projects, and examined the supporting literature of online systems and publications accompanying them, and discuss their design choices in the context of previous guidelines on building successful online communities from the greater HCI and CSCW literature. Our study documents significant variations in the design of specific features, including rewards, performance feedback, goal setting, and seeking and managing contributions. It further identifies areas where CS projects are often incomplete or disregard best practises. We describe the online citizen science design space as a whole, to enable further research and innovation in this area.

Summary of Contributions

In this work we carried out a survey of online community design features in 48 citizen science projects, drawn from a large-scale meta study of over 150 scientific publications. To the best of our knowledge, this is the first work of this scale where a comprehensive set of CS project features are identified, extracted, and systematically analysed from the angle of online communities. Our research provides insights into the impact of those features on success of the underlying CS projects and specifically identifies community features as an essential component. Such insights are of particular benefit to those wishing to build successful CS projects with collaborative community aspects, highlighting key design considerations for overall project success.

BACKGROUND

In the past ten years there has been a rising interest in crowdsourcing approaches to scientific enquiry and experimentation. Such endeavours range from participatory sensing, to human computation projects in which volunteers collect, curate, and analyse scientific data. Our meta-analysis focuses on the extent to which existing citizen science projects have been able to build successful communities that contribute to their scientific aims. As argued in [59], the emergence of these communities has proven critical for the success of citizen science initiatives. Not only have such communities played a crucial role in several scientific discoveries, but they have also often helped new projects and sub-communities to form and grow. The section is divided into two parts: an introduction to citizen science, followed by an overview of previous work on online community design frameworks.

Online citizen science

Online citizen science draws upon theories and practice from several areas that have long been studied in HCI and CSCW, including human computation and crowdsourcing, online communities, and gamification. They have often been characterised as crowdsourced science [37], where professional scientists seek the help of large numbers of people to contribute to scientific research [14]. In its most common instance, a project would seek the help of volunteers to take on ‘microtasks’, collecting, curating, annotating, and analysing scientific data at a level that does not require specific knowledge or domain expertise [59]. Done effectively, such microtasks are dispatched and results are validated and aggregated in ways that allow project scientists to process large amounts of data accurately and at high speed [33]. One of the most prominent CS projects is *Galaxy Zoo*, which attracted more than 50,000 astronomy enthusiasts who classified hundreds of thousands of galaxies in just a few weeks. Such a task would have been extremely expensive and time-consuming if done by scientists alone or with the help of state-of-the-art object recognition software [27].

In a classification project such as *Galaxy Zoo*, participants are presented with an entity, which can be an image, graphs, an audio file, or a video, and then asked questions about that entity. For instance, users may be asked to identify features, map these features, catalogue entities, transcribe, or complete other microtasks [12]. Figure 1 shows a user interface from *Snapshot Serengeti*, from the *Zooniverse* platform. Here, the large image on the left corresponds to the entity which must be classified. The menu on the right lists the animal species to choose from. Microtasks may take various other forms besides classification. For example, in *eBird* volunteers submit observations of bird distributions in real-time. Some projects are minimally designed online spaces for people to upload or share their data, while others have more complex interfaces, sometimes in the form of games [8].

Despite involving large numbers of people, citizen science has historically targeted individual participants rather than groups or a larger community. It is thus uncommon for participants to be made aware of each others’ contributions, and microtasks are typically designed to be solved independently [33, 61]. Success is often measured in these terms, with metrics such as the number of registered users or the time it took to complete a certain goal being popularly stated measures[1]. Responses for each entity are typically aggregated and compared as a means of validation [20]. In the event that there is significant disparity between responses, an entity will be presented to a greater number of participants to gather further evidence.

Citizen Science projects vary in aim between those which seek to complete scientific research, with a focus on investigation, and those which seek to engage volunteers in science through education, training and raising awareness of issues such as conservation. This variation manifests in a number of characteristics, including the tasks requested of users, the technologies used in data collection and analysis and the settings, physical and virtual, in which these processes take place [62]. Just 16% of papers surveyed by Kullenberg and Kasperowski were found to have a scientific research output, reflecting the importance of these alternative educational, engagement goals [23]. At the same time, citizen science projects are associated with a relatively slow rate of publication [56].

Similarly, projects differ in the stages of the scientific process in which citizen volunteers are invited to participate. At their simplest, citizen science projects take the form of ’volunteered computing’ or use sensors, with volunteers used solely as a means to access processing power or distributed technology [13]. A more extreme form of citizen science is ’collaborative science’, an approach where volunteers assist in defining a research problem, gathering and analysing data and may even design studies, draw conclusions and disseminate findings [63, 13].

[1] Zooniverse blog, “Measuring success in citizen science projects, part 1: methods” – http://blog.zooniverse.org/2015/08/24/measuring-success-in-citizen-science-projects-part-1-methods/ [Accessed: 27 May 2016].

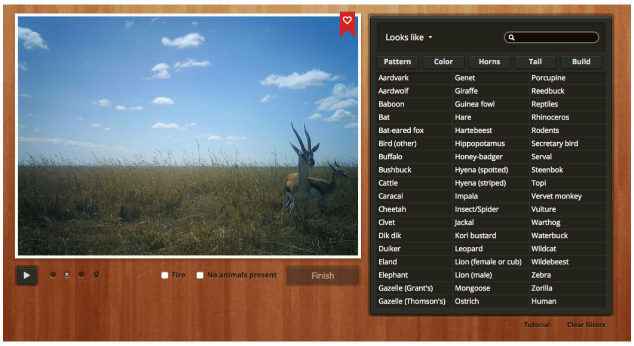


Figure 1. The classification interface in *Zooniverse*'s *Snapshot Serengeti* project asks participants to select the species of animals present within images from an extensive list.

As the field of CS evolved, it became clear that introducing mechanisms for collaboration would be beneficial for the performance of the individual contributors and their long-term engagement with the project [59]. These mechanisms have taken diverse forms such as discussion forums [40], instant messaging services [19], or custom-built community spaces (for example, *Zooniverse*'s *Talk*). Such spaces allow for communication both between participants and between participants and the scientists responsible for each project. Since users are able to directly interact with one another, scientists are able to rely on the community to answer user queries, or to direct such queries to scientists when appropriate, rather than having to respond to queries on an individual basis [59]. There is evidence of communities gaining scientific knowledge through the use of discussion features [29]. Furthermore, several important serendipitous discoveries have grown out of these interactions: Hanny's Voorwerp and Green Peas in the context of *Galaxy Zoo*, [41, 5], the Circumbinary planet PH1b from *Planet Hunters* [46], and the new variety of nebula in *The Milky Way* project [18]. These discoveries have themselves resulted in publications in academic journals, written in conjunction with volunteers.

Online community design

Online communities are environments in which people gather to work towards common goals, socialise, share knowledge and resources, or communicate [22]. Such communities vary in size and may take diverse forms, although the majority of online communities take the form of textual discussion forums or email groups [24, 22]. Online citizen science projects are examples of online communities. A large community of volunteers must come together to enable completion of scientific goals, which often requires classifying tens of thousands of entities [59]. Online community design is therefore an important consideration when designing successful CS projects.

Online community design has been studied using various frameworks and analytical methods. Ren, Kraut, and Kiesler analysed theories of commitment with regard to online communities in order to identify design decisions which may lead to greater commitment [42]. The authors demonstrated that specific design decisions such as constraining or encouraging discussion among the community can influence group forming and commitment between community members and so influence the form that community participation takes.

Preece [38] explored the dimensions of sociability and usability in an effort to explore the concept of success in online communities. The resulting framework considers facets of sociability: volume of participation, reciprocity in contributions and benefits, quality of contributions, and participant behaviour. Usability dimensions are also key components of the framework: ease of use, speed of learning, measures of productivity, and user retention. These dimensions are of particular interest given our objective to identify features associated with online community success in CS projects.

Similarly, Iriberri, and Leroy analysed online communities within the framework of information systems life cycles to further explore the concept of success in online communities [15]. Dimensions evolve throughout the cyclical framework, from conception and purpose, to ensuring security and reliability during the creation process. Quality assurance and encouraging interaction become important during the growth phase, while mature communities must further focus on rewarding and encouraging interactions through events. Such a framework demonstrates the evolving nature of success and highlights the importance of early design decisions on later project outcomes.

Kraut et al. studied a diverse body of communities, ranging from crowdsourcing efforts such as *CAPTCHA* and *Mechanical Turk* to MMOs such as *World of Warcraft* [22]. Each of these communities was analysed based on empirical observations informed by key theories from the social sciences. They devised a total of 175 design claims, across five areas: encouraging contribution, encouraging commitment, regulating behaviour, dealing with newcomers and starting new communities. These design claims provide evidence-based guidelines for assessing online community success.

With regard to gamification, Mekler et al explored the impact of points and contextual framing of tasks on volunteers' intrinsic motivations and performance in an image annotation task [31]. Points were found to increase the quantity of tags generated, while framing had no significant effect on quantity. A combination of points and contextual framing was shown to have a significant effect on the time spent per tag, compared to just points or framing alone. Framing was associated with an increase in tag quality, an effect which was not seen with points. These findings suggest interface features can impact participant engagement and the effort expended by participants.

DATA AND METHODS

Our goal was to identify design principles and guidelines for improving CS projects by observing the interplay between current design decisions and success within project research outcomes. To this end, we consulted both the project platforms and research output from projects and the wider online citizen science literature. In order to do so, we began our analysis by identifying a selection of projects and documentation through a review of contemporary CS literature. Drawing from the success metrics identified during the literature review, we developed a framework of themes and success criteria affected by such themes within online community literature. This framework was used to conduct a survey of the selected projects, identifying mechanisms and affordances pertaining to each theme through a structured walkthrough process. The evidence basis determined through our literature review was further used to analyse the results of the survey process, to provide evidence for our findings from the wider literature and to distill design implications for further research moving forward.

Selecting Projects and Project Documentation

Initially, projects and project documentation were selected through a literature review designed to identify suitable CS projects, as well as potential issues pertaining to online communities. A search was conducted of four online repositories, selected to identify publications from a broad range of research disciplines (see Table 1). This search resulted in a total of 886 publications. The title and abstract of each were assessed to remove duplicates and publications deemed to be irrelevant (for example, offline CS systems). 152 publications were thus selected for further use.

Each publication was first analysed to identify potential projects for sampling. Projects were considered for inclusion if they contained at least one community feature (forum/wiki/discussion board/IM chat) and if it remained possible to register and contribute to the project. We excluded social media features due to the lack of input which social media users have on platform design and the difficulties in identifying official project social media channels.

After completing this process, we found that the majority of projects were no longer available, a factor which we partly attribute to the prolonged time it takes for citizen science data to reach publication (see, for example, [56]). Aware of the success that participants in citizen science systems have had in spreading awareness of systems (see, for example, [59]), we supplemented our sample with projects listed on *Wikipedia*, which features a well-maintained list of citizen science projects[2]. To ensure the validity of the information gathered, we visited each project URL to confirm the existence and suitability of the project.

We identified 136 projects in total, from which we selected 48 projects for further analysis according to the inclusion criteria. This list of projects was not intended to be exhaustive; while we made efforts to ensure the inclusion of a range of project microtask types and disciplines, the projects selected for inclusion are likely to be more popular than average, having drawn academic or volunteer interest. This was a deliberate choice, resulting form our desire to assess design decisions in projects making successful use of design features. Although we included multiple projects from the *Zooniverse* platform, we note that specific design decisions vary between these projects and we thus consider these projects individually. The full list of surveyed projects can be seen in the appendix.

Having identified a suitable sample of projects, we returned to our sample of literature and conducted a systematic review of each publication. We first removed publications deemed unsuitable for further use. This predominantly consisted of those publications which served only as data-releases of projects for which other, more informative publications were available. We also removed publications which referred only to projects deemed unsuitable for inclusion within our study. This generated a smaller sample of 115 items of literature. Each publication was reviewed in more detail to identify relevance to the surveyed projects, relevance to each of the four themes utilised within our online community framework and to determine success metrics and evaluation methods utilised within the literature.

[2]Wikipedia, "List of citizen science projects" – https://en.wikipedia.org/wiki/list_of_citizen_science_projects [Last Accessed: 27 May 2016]

Repository	Search terms	Results
JSTOR	ab:("online" + "citizen science") OR ab:("digital" + "citizen science") OR ab:("virtual" + "citizen science")	9
Google Scholar	"online citizen science" OR "virtual citizen science" OR "digital citizen science"	424
Scopus	TITLE-ABS-KEY ("online" + "citizen science") OR TITLE-ABS-KEY ("digital" + "citizen science") OR TITLE-ABS-KEY ("virtual" + "citizen science")	242
Web of Science	TOPIC:("online" + "citizen science") OR TOPIC:("digital" + "citizen science") OR Topic:("virtual" + "citizen science")	211

Table 1. Repositories and search terms used for the literature review.

Online Community Framework

To identify the extent to which projects adhered to online community design recommendations, we synthesised a framework of recommendations based on existing literature. We drew from the work of Kraut et al's *Building Successful Communities* [22] due to its application to a range of online communities and because it provides specific design claims for social and technical characteristics of successful online communities. These were considered alongside the frameworks described by Nov et al [35], Iriberri and Leroy [15], and Preece [38].

We first extracted design recommendations from the literature, identifying over 200 unique design recommendations. To ensure relevance to citizen science, we selected only recommendations which relate to ensuring high quantity and quality of contributions, as identified within the originating frameworks. Further, we selected only recommendations which we could observe from the systems alone; recommendations which would require consultation with participants were deemed beyond the scope of this project given the large number of systems involved.

Recommendations were finally grouped into a total of four broad themes:

- *Task visibility* – the ease with which participants can see and select microtasks and discussions requiring completion.
- *Goals* – the provision of challenges and targets for participants to achieve.
- *Feedback* – mechanisms for informing participants of the quantity or quality of submissions.
- *Rewards* – tangible or intangible awards given to participants for making contributions or achieving goals.

Structured Walkthroughs

To assess these themes within each of the selected projects, we conducted our survey by utilising a structured walkthrough-based approach. For each project, two researchers registered as participants and completed approximately ten classifications (or for data collection projects, assessed existing contributions), as well as observing community interactions. This was a four step process: Initially, each of the researchers separately registered and completed 10 classifications or other forms of contribution (evaluating existing contributions in data collection projects) within each of the 48 projects analysed within the study. During this initial step, each researcher produced a list of affordances, mechanisms and characteristics observed across all 48 projects. Following this, both researchers consulted with one another and compared the two lists, discussing points of contention and disagreement between the two lists, referring to the specific projects involved, in order to produce a common, unified list of mechanisms agreed upon by both researchers. Each researcher then separately returned to each of the projects, completing further contributions where necessary, in order to survey the number of times each of the mechanisms within the list was utilised within the sampled projects. Upon completion of this survey, the two researchers convened again to correct errors, discuss disagreement and produce a final list detailing the observed mechanisms and the number of occurrences of the mechanism across the projects. In order to ensure accuracy and to prevent possible issues with the structured walkthrough approach, this list was then compared with evidence drawn from the literature review process, using relevant publications where available for each project, as well as project blogs and news-feeds.

RESULTS AND ANALYSIS

In this section we report on our cross-sectional analysis of 48 citizen science projects by using the structured walkthrough, organised by the four themes: task visibility, goals, feedback, and rewards. For each theme, we identified a list of commonly occurring mechanisms, ordered by frequency within the sampled projects.

Task Visibility

Many of the identified mechanisms facilitated making tasks and discussions visible to users, such that volunteers could identify entities which require classification or discussion which requires participation (see Table 2). What follows is a synthesis of affordances based on our observed use of these mechanisms.

Automatic entity selection. In terms of microtask contributions, a significant proportion of platforms automatically selected entities for volunteers at the start of each session, providing little or no indication of how these entities were chosen behind the scenes. This way of assigning microtasks allows the science team to control the number of times each entity is classified by the crowd, while also ensuring completion. The *Zooniverse* platform uses an algorithm to control the number of classifications that each entity receives, although, in practice, entities may receive a greater number of classifications if the number of participants outweighs the number of available entities [50, 53]. *InstantWild* limits the available number of images for classification, with eight entities available for classification at any one time, while *EteRNA* allows contributions to each round of its *Cloud Laboratory* for a limited period of time only. Whilst still automated in task-selection, *EteRNA* allowed participants to select from all available entities by solving puzzles, with the ability to filter puzzles by recency, rewards, number of prior completions, and length. Similarly, *FoldIt* offered participants the opportunity to select puzzles, with several pre-existing groupings offered.

Mechanism	# of projects
Notification of most recent activity	48
Free selection of discussion threads	47
Automatic Assignment of Entities	44
Sticky/pin function	43
Entity availability limited by classifications received	41
Follow function (by thread)	34
Completion percentage (by collection)	7
Dedicated area for entities in need of input	5
Customisable discussion feed	2
Entity availability limited by total number of entities	1
Task available for limited time	1
Follow function (by entity)	1

Table 2. Mechanisms which support task visibility.

User-task selection. While automatic selection discourages large amounts of activity around particular entities, community-specific features may encourage a disproportionate amount of activity around a given task. Community-specific features tend to be facilitated with discussion boards, offering specific threads on a popular topic, particularly topics which have recently received attention from other volunteers. Participants are able to select threads and discussions freely, regardless of the number of discussions surrounding that entity, or time since the entity was uploaded. However, this does not extend to the microtask interface, which prevented the selection of specific entities in all but 4 of the most gamified projects such as *EteRNA* and *FoldIt*. In *EyeWire*, entity selection is a specific privilege, offered to a small number of participants as a reward for contributions of particular value (see *Rewards* for more detail).

Drawing attention. CS projects function by drawing volunteers' attention to entities requiring additional work, or to previously completed work as a learning experience. This was particularly common in community features, where threads or comments could be made clearly visible through the use of sticky or pin functions, causing these threads to remain at the top of any lists of discussion threads. A similar mechanism was used in the task area of five projects, where a dedicated, clearly visible area was reserved to draw participant attention to entities in need of work. In community features, attention was almost always drawn to the most recent discussion contributions, as a proxy for those in need of contribution.

One area lacking from task visibility is the opportunity for volunteers to easily select entities which appeal to them. The 48 projects surveyed all offered participants relatively low levels of autonomy with regard to task visibility and task selection. In some cases projects would allow volunteers to select specific collections, such as logs from a given ship in *Old Weather*, or images of specific kingdoms and/or classes (e.g., fungi, birds and insects) in the *Notes from Nature* project. However, the difference between collections is largely thematic. In these cases, the burden of choice was placed on the user, rather than recommending collections based on participants' previous behaviour.

Mechanism	# of projects
Classification challenges	18
Competitions	14
Opportunity for Rare Discoveries	6
Meta challenges (fundraising, attracting attention)	5
Survey (user voting for entity naming, etc.)	4

Table 3. Mechanisms which support goals.

Phylo and *FoldIt* allow volunteers to complete projects aimed at understanding specific diseases, rather than making use of the random assignment function, but this was again largely based on participants' choices rather than specific recommendations.

Goals

One effective method of motivating contributions identified within the literature concerns is to assign volunteers achievable goals. We identified a number of goal-setting mechanisms in various forms, as shown in Table 3, including the use of challenges and competitions to achieve a collective benefit, as well as individual benefits. Below are three forms of goals which these mechanisms afford.

Project-completeness goals. The most common form of goals observed, surrounded task completion and, in particular, the number of contributions received, with volunteers asked to increase their level of participation to meet these goals through *classification challenges*, coupled with *competitions* in 14 projects, predominantly on a temporary basis. The *Planet Hunters*, *SpaceWarps*, *Planet Four*, and *Higgs Hunters* projects all made use of three day challenges, asking volunteers to complete as many contributions as possible, to coincide with the BBC's *Stargazing Live* broadcast. These temporary goals lead to brief periods of extremely high rates of classification – the *SpaceWarps* project generated 6.5 million classifications in just 3 days, with a peak of 2,000 classifications per second [49]. After the completion of the challenge, contribution rates fall sharply – after 2 years, the three day challenge still accounted for the majority of contributions to the Planet Four project [45]. In the same manner, progress bars and completion counters are used to indicate the state of a project. The specific size and nature of a collection varies between projects. *Snapshot Serengeti* and *Verb Corner* divide entities into collections based on the period of time over which images were gathered, requiring completion of one season before another can be released. *Notes from Nature* runs concurrent collections, divided based on the focus of entities (e.g., plant, bird, insect) and thus the fields required for transcription. Other collections are more thematic – *Old Weather* divides log book pages into collections based on the ship from which the log book was taken. While for the most part these goals had specific deadlines, those mechanisms which served dual purposes lacked deadlines. Collection completion percentages, for example, function as goals and as a means of making tasks visible. Collection completion was not tied to specific deadlines – collections remained available and accessible until they generated sufficient numbers of classifications, at which point they were removed.

Milestone-driven goals. In some cases, challenges did not correspond to the completion of a collection, but to a set level of contribution, in the forms of milestones such as 'one million classifications'. Moon Mappers ran the 'Million Crater Challenge': setting a goal to achieve one million crater classifications across all participants between April the 20th and May the 5th, 2012. Major milestones offered rewards to individuals, as a further level of benefit. Although volunteers successfully completed over 100,000 classifications within the time-limit set for the challenge, the goal ultimately proved too challenging and it was not until October that the goal was finally reached. *EyeWire* has similarly hosted a number of month-long classification challenges. Participants are assigned a team at the beginning of the month and must score as many points as possible for their team throughout the month. *EyeWire* has also offered volunteers the chance to take place in a number of competitions at an individual level, with the aim of achieving the highest level of accuracy or the greatest number of classifications. Such tasks may involve scoring the highest number of points, making the most classifications or achieving the highest level of accuracy over the course of a week. Prizes are awarded to the winners (see Rewards below).

Community-based goals. The existence of goals aimed at community-feature participation is rare, with such goals almost exclusively taking the form of meta-challenges, aimed at aiding in the administration of a project or public awareness. These community-based goals were those most likely to affect the wider public, outside of the community of project participants. *Planet Hunters*, a project which aims to discover new planets, runs occasional competitions to allow volunteers to name new planetary candidates. Participation in the competition is entirely through community features, such as forums and survey forms. A similar community challenge, the *Snapshot Serengeti* 'Serengeti Selfies' campaign, aims to raise funds by asking participants to identify images of animals which appeared similar to photographic self-portraits in the 'selfie' style, for publication. Participants are asked to use the hashtag #selfie to identify such entities within the *Snapshot Serengeti* talk pages. This does not require engagement with the task-interface, though unlike Planet Hunters, there is nothing preventing the identification of such images through the task interface and thus, does not exclude participation through task completion. When the *Snapshot Serengeti* project was faced with a reduction in funding, the team began a crowd-funding campaign. To publicise the campaign, a concurrent campaign was launched, where, while completing classifications, participants were encouraged to find amusing or interesting images and caption them, before sharing them via social media. The campaign successfully raised $36,324 and, at its peak, attracted 4,500 unique users to the *Snapshot Serengeti* project [21].

Feedback

In the systems observed, feedback serves a dual purpose, helping to ensure the validity of results, while also engaging volunteers through learning. Feedback provision mechanisms can broadly be divided into two groups: task-contingent, related to the number of completed tasks, or performance-contingent, related to the quality of contributions received [22]. Furthermore, feedback may be quantitative and systematic, through the task interface or discussion-based, through community features such as forum discussion.

Mechanism	# of projects
Task-contingent feedback	29
Performance-contingent feedback	14
Performance-contingent feedback as numerical score	11
Gold Standards for performance-contingent feedback provision	7
Progress-bars for task-contingent feedback	7
Volunteer testing	5
Majority opinion-based performance-contingent feedback	4
Comments from science team	1

Table 4. Mechanisms which provide participants with feedback on individual contributions and overall project progress.

Task-driven feedback. Task-contingent feedback was relatively common among the projects surveyed. The *SpaceWarps* project interface provided a counter which displays the number of images a user has viewed, as well as the number of potential gravitational lenses discovered. Such feedback may also be provided through comparison with other community members; *Herbaria@home* tracks all user contributions in a leaderboard, *Old Weather*'s rank function divides leaderboards into bands (ranks) where participants may progress by completing more classifications, *EteRNA*'s point system is a relatively unique feedback mechanism, with participants gaining points based on the difficulty of a puzzle, rather than solely the quality of their response. [54].

Performance-driven feedback. Explicit performance-driven feedback mechanisms were rarer among the projects surveyed. One relatively unique form of performance-contingent feedback occurred in *Phylo*; participants must pair nucleotides to build DNA sequences, with the task encouraging participants to match similarly coloured blocks. Participants are assigned stats based on how they perform in the task, which is then used to construct global leaderboards. In *EyeWire*, players received feedback as a point score, this score varied according to a number of factors, including the difficulty of a classification and the extent to which the classification differed from the average classification received [43]. Within the FoldIt project, puzzles were difficult and the research conducted was often complex for participants to understand. Volunteers therefore relied on these point scores to understand how their performance matches with what is expected from them and whether they were giving a useful or correct answer [11]. In *InstantWild*, participants saw an anonymised summary of the classifications received from other participants for a given entity. One inherent vulnerability in such a majority-based feedback mechanism as cited by the *EyeWire* team, however, is the danger of the majority opinion being incorrect, introducing the possibility that *EyeWire* classifications which are more correct than the majority will earn fewer points [43]. Performance-driven feedback has been shown to be highly effective in the EteRNA project, where participants modified their approach to puzzles according to results from laboratory experiments derived from the most effective submissions as judged by project scientists [9].

Performance-related feedback can also be combined with Gold Standard data; entities for which the 'correct' response is already known. By asking participants to classify these images, project administrators can compare responses to the predetermined expert response in order to provide feedback to participants. These gold standards may be determined in a number of ways; *Stardust@home* makes use of an algorithm to classify gold standards. These gold standards, known as 'power movies' were periodically shown to participants, who received feedback by email (or within a report within the classification interface) detailing the number of power movies found, the number missed and a numerical score to describe their performance. *CosmoQuest* projects took a similar approach, using expert classified gold standards, with participants given a numerical score for their performance. *SpaceWarps* made use of simulations, with textual pop-ups indicating correct/incorrect responses.

In 5 projects, performance-related feedback took the form of a test. *Stardust@home* prospective participants must pass a test to contribute to the project, by proving their ability to identify interstellar dust particles within entities. Participants are informed of correct and incorrect answers after completing the test. Each of the 4 *CosmoQuest* projects also tested volunteers, interspersing classifications with small tests, which provide feedback to participants. Unlike in *Stardust@home*, participants do not initially need to pass a test to contribute, but the tests encourage participants to repeat the tutorial process if a low score is earned.

Community feedback. In addition to feedback offered by site administrators, each of the surveyed projects offered participants the chance to provide written feedback in their community postings. A subset of projects explicitly encouraged this form of feedback; *Stardust@home* and *Herbaria@home* both featured long-running, 'stickied' forum threads with the purpose of allowing users to give and receive feedback on classifications. Even outside of these projects, feedback appears to be a common usage of discussion features, with *Zooniverse's Talk* interface allowing users to tag images with classifications, as a means of receiving feedback from other participants. The extent to which such feedback occurs is arguable, however – studies have demonstrated that in excess of 90% of discussion comments on *Zooniverse's Talk* platform have gathered no replies [29]. *Stardust@home* users received direct feedback from members of the *Stardust@home* science team. An area of the feedback page for each user contained space for comments from the science team in response to potential interstellar dust candidates. As with citizen-sourced feedback, however, this feedback remains rare. As of August 2016, only 88 potential dust candidates had been discovered since the project began and only 12 of these have feedback from members of the science team [55].

Measuring feedback. Task-contingent, quantity-based feedback was delivered exclusively in a quantitative form, with projects such as *SpaceWarps* and *Herbaria@home* keeping track of classifications in the form of a numerical score. Performance-contingent feedback, however, was offered in both quantitative and qualitative forms. Systems such as *Stardust@home*, *EyeWire* and *VerbCorner* kept track of numerical scores for participants as a form of feedback. Similarly, *CosmoQuest* participants received numerical scores after completing gold standard classifications, as a form of feedback. *SpaceWarps*, conversely, made use of written, qualitative pop-ups to provide feedback to participants on gold standard classifications. Where feedback was provided quantitatively, it took the form of point scores, overlapping with reward mechanisms. However, in two of these projects (*EyeWire* and *VerbCorner*), the method used for calculating scores was somewhat hidden from participants, to prevent efforts to game the system. This in turn makes ascertaining specific feedback, such as feedback on accuracy, relatively difficult, as accuracy-based scoring could be separated from other factors. In *EyeWire*, for example, accuracy was expressed by removing points from a user's score based on any perceived inaccuracy [43]. Participants did not receive a prompt indicating the number of points removed and the maximum score a user could receive varied between classification entities. As a result, while feedback calculation is systematic, the manner in which it is expressed to participants is less so.

Rewards

Contemporary literature argues that rewards encourage people to provide contributions. As Table 5 describes, across the surveyed projects, rewards can be supported through a variety of mechanisms. Rewards could be awarded based on the quantity of responses, as task-contingent rewards, or on the quality of responses, as performance-contingent rewards.

Rewards within the surveyed projects took one of three key forms. Status rewards function by increasing a user's reputation, elevating that user's status by making other participants aware of his/her achievements. Privilege rewards allowed volunteers to access additional tools or task types, which other participants with lower levels of participation could not access. Physical rewards refer to prizes such as project merchandise.

Status rewards. The provision of reputation rewards was common across many projects. *EyeWire* participants who consistently performed at a high level of accuracy may receive a promotion to the role of *scout*. Scouts are identified within the *EyeWire* task interface by turquoise user names and through the word scout within their user profile. Roles were also granted within community features. Users of *Zooniverse*'s *Talk* were on occasion selected to serve as moderators, receiving the tag 'moderator' next to their posts. Leaderboards also served as a status reward, particularly for users in high positions. Leaderboard calculations varied between systems. Some leaderboard calculations were solely task-contingent, such as *Herbaria@home*, where users were ranked solely by the number of classifications they have made. Others were performance-contingent, as in the case of *EyeWire* and *Stardust@home*, where user rankings used points calculated based on accuracy. Leaderboards varied based on the time-scale over which ranks were calculated. *Herbaria@home* rankings described user contributions over the entire life of the project, while *Stardust@home* rankings were divided into seasons. *EyeWire* rankings were divided into three short-term categories: "today", "this week", and "this month.". *Old Weather* featured a rank system which served similarly to task-contingent leaderboards. As participants completed classifications within a given collection of entities, their number of classifications was compared with that of other users.

Mechanism	# of projects
Status rewards: Titles/Roles	41
Status rewards: Leaderboards	11
Points	11
Task-contingent rewards	11
Public announcement of achievements	7
Status rewards: Achievements/badges	5
Physical rewards	4
Unrevealed reward calculation factors	2
Privilege rewards: Additional tasks	2
Privilege rewards: Entity selection	1

Table 5. Mechanisms which reward participants and incentivise contributions.

We found a number of platforms used badges/achievements to encourage contribution. These were awarded to participants for achieving particular goals within systems. Each collection within *Notes from Nature*, for example, had associated badges. These badges were entirely task-contingent; participants earned badges by completing a given number of classifications. Similarly, the *EyeWire* forum made use of task-contingent badges, with participants receiving badges for completing specific forum activities such as creating a post; editing a post or liking a post. The *EyeWire* classification task featured similar but more complex achievements. Achievement requirements are not advertised to participants and instead users discover achievements as they classify. Requirements may be task-contingent (such as completing a tutorial) or performance-contingent (such as earning a certain number of points or a certain level of accuracy.)

Privilege-driven rewards. Privilege rewards were often linked to status rewards; participants receive privileges either as an indication of their status within the community, or alongside a status reward. *EyeWire* roles, for example, had accompanying privileges, unlocked when users were promoted to a given role. Scouts were able to inspect any entity within a collection, rather than having to rely on automatic assignment. Similarly, moderators in *Zooniverse* talk were able to moderate discussions by carrying out activities such as deleting posts, something ordinary users are unable to do. In all cases, roles granted privileges in community-based activities. *EyeWire* role recipients received special indications in the live IM-style chat function and special chat tools. Similarly, *Zooniverse*'s moderators received special tools and capacities to moderate discussions, while receiving an indication of their moderator status. Other privilege rewards served independently of status rewards. The *VerbCorner* project allows users who have completed a certain number of contributions to unlock additional task types.

Physical rewards. Physical rewards were particularly rare among the projects surveyed. Only four systems offered physical rewards, in the form of physical prizes. Rather than using physical rewards for regular forms of participation, these rewards were only available temporarily and only for a very small number of users, who were responsible for particular achievements, or outright winners of a challenge. *Moon Mappers* offered up to 20 physical rewards during its Million Crater Challenge, for the user responsible for each 100,000th classification and ten randomly selected users. *EyeWire*, meanwhile, ran a week long *Camp EyeWire* event, with challenges based on accuracy and challenges unrelated to the project (such as a trivia quiz). Each challenge awarded both physical rewards (in the form of *EyeWire*-related merchandise) and bonus points to the winners and runner ups.

Reward exploitation. One issue posed by rewards is the danger of exploitation. With the introduction of rewards, some users shift their focus to achieve maximum rewards with minimum effort; completing tasks inaccurately or otherwise gaming the system. Certain projects aimed to counteract this by rewarding users on an unpredictable schedule and obfuscating reward criteria; a method employed by both *EyeWire* and *VerbCorner*. The points awarded to *EyeWire* users were calculated based on a number of dimensions, from accuracy to time. However, while spending more time on a classification was associated with a higher score, time-related points were capped at an unspecified value, making it difficult for users to accumulate points from delaying classifications. Further, an explanation of how points are calculated was kept relatively vague [43]. Similarly *VerbCorner* users received bonus points while contributing, but the requirements for bonus points were not publicised [2]. *LandscapeWatch Southampton* features task-contingent rewards in the form of points. However, point scores are non-linear and users are not informed how points are calculated.

DISCUSSION

In this section, we discuss two areas which are relevant to the ongoing debates in citizen science and online community literature as identified during the literature review process: platform design for microtask- and community-orientated support, and factors to measure successful citizen science projects.

Microtask-Orientated vs Community-Orientated Features

Based on our structured walkthrough of 48 citizen science projects, we have observed common features and design patterns across the projects. By framing our analysis around 4 themes, namely: task visibility, goals, rewards, and feedback, we observed a spectrum of features pertaining to the management of two critical components; the task, and the community.

We have structured our results around four separate themes for purposes of clarity, and to reflect existing online community framework literature. It is however clear, that these four themes are highly coupled with each other. Each component is integral not only to ensuring user engagement, but also to improving and maintaining consistency and accurate results. Approaches to goal-setting and reward provision are closely linked to the way in which microtasks are presented to participants. Furthermore, it is these goals and rewards which determine the type of feedback which should be provided to users, as well as the form this feedback should take.

In light of our observations and analysis, our findings suggest a closer connection between microtask and community features than has previously been considered. While rewards predominantly result from engagement with and completion of project microtasks, their value predominantly results from community prestige: leaderboards, badges and titles, for example. In some cases, titles and roles were granted, conferring community-related privileges (for example, the opportunity to moderate conversations, or to interact with users requiring guidance), based on microtask contributions. Goals attached to deadlines, for example, are attached to the completion of classifications and even meta challenges, aimed to maintain projects through community action may result in increased rates of microtask completion. Furthermore, interaction with fellow community members fulfils an important role, particularly in those projects which lack other feedback mechanisms.

In the systems we have studied, the design of the microtask, and the features which provide the interaction layer are vastly different. As the analysis has revealed, the various systems provide different levels of interaction with the microtask component of the citizen science project. Automatic microtask selection was a dominant feature in many of the projects, despite existing literature suggesting that participant performance may be increased when manual selection of microtasks are possible [22].

With regard to automatic selection, by identifying user strengths, it is proposed that projects could be made more efficient, with users shown entities which are related to these strengths, rather than random entities. Similarly, by understanding appealing classification entities, users could be offered a greater proportion of entities which interest them, with an aim of ensuring they remain motivated throughout the life of the project. Within the *Zooniverse* platform, there have been proposals and experiments conducted to implement features surrounding this area, along 3 dimensions: user ability, user interest and maximising user motivation [26], however due to the complexity of microtask-assignment, the algorithm used by the *Zooniverse* makes use of random assignment of entities [50]. Furthermore, there is a trade-off here: allowing participants to select their own tasks has the danger that classifications become disproportionate, and certain resources are never selected. An alternative to this would be to design the task selection as a mix of automation and manual selection. EteRNA is a good example of such a workflow, where participants can select tasks from a pool of pre-assigned resources.

As our analysis revealed, community-supported task selection was not favoured, and was not used (as the main mechanism) by any of the projects. Existing studies have shown projects which have strong communities tend to be more successful, generating more classifications, and taking less time to complete project goals[29, 59]. Whilst we found projects such as *EyeWire* enabling users and admins to recommend specific tasks to work on individually and via teamplay mode (this is not the main approach, random assignment is dominant), there is very little effort towards letting the crowd assign the tasks, or at least recommending tasks for participants to complete. We also see this highly relevant to the lack of user-operated tools, which provide control and power to the user with respects to completing tasks, or more generally, interacting with the system. In studies of non citizen science online communities, user tools have been shown to improve overall success of the platform [22]. For citizen science, tools to enable users to monitor and track their work could enable new modes of operation, facilitating the discovery of tasks which may be relevant to a player, clustering like-minded players based on their skills and past classification history, or to help the serendipitous discovery of scientific knowledge [59]. Moreover, such tools can also be coupled with the community, allowing collaborative task workflows to emerge.

Closely related to the task visibility is the implementation and integration of feedback mechanisms to provide participants with guidance and reassurance on their contributions. As our analysis reveals, feedback mechanisms were not a common feature, despite users requesting such features (e.g. *Galazy Zoo* users [39]). Roy et al., [44] note the importance of feedback mechanisms for maintaining user engagement and motivation in citizen science initiatives. However, providing feedback is complex and often difficult as many of the tasks do not have a correct answer [60]. Overcoming the difficulty of providing timely and meaningful feedback can be achieved via a number of methods beyond automated methods; as Kraut et al. [22] describe, community-driven, discussion based feedback is often a suitable method when it is difficult to obtain repeatable quantitative systematic feedback. As a number of projects have shown, community-driven feedback can yield highly valuable results, such as unanticipated scientific findings [5]), or in projects such as *EyeWire*, where features like the real-time chat interface, have helped harness the expertise and knowledge of long-term members to encourage newcomers, and facilitate the crowdsourced answering of players.

Goals and rewards are also major components to be considered for a citizen science project. These depend on the type of task, and more importantly, the decisions made at the initial stage of designing the platform. Considering the projects reviewed, the use of gamification elements, such as leaderboards, points, badges, and status are important design decisions that have to be made, which have implications for the future of the projects marketing, community management, and maintenance (socially and technically). These decisions also have implications on the community component of a project; for instance, in *EyeWire* [58, 57], competitive elements encourage participation. Such phenomenon has been observed elsewhere in other online community platforms [64], where up to three times as many contributions were found when suitable goals and rewards were used. Given the strong intrinsic motivations to participate [41, 57], designing these features with a strong emphasis towards community engagement is beneficial to a projects success. Even where gamification and competition is integral to the design of the platform (c.f. *EyeWire*), participants use their elevated-privileges to further support their fellow participants.

Success factors for Citizen Science

Citizen science is a diverse and growing field, with a range of task types and scientific goals. Perhaps because of this, there is currently no universally accepted set of criteria for defining project success. Those studies which have attempted to define such criteria generally apply to a given project, platform or context, such as the work of Cox et al. which utilises *Zooniverse* as a basis for success criteria [7]. To determine success metrics, we identified common measures discussed within CS literature, as identified during our literature review, which may serve as a basis for such success metrics. While there is no commonly accepted framework, we believe these measures describe common aims across citizen science projects.

Engagement (i) – Number of people reached

One commonly cited statistic within CS literature concerns the number of users which a given project has engaged and who have contributed to the project, both through microtask and community contributions. Within our survey, *goals* and *rewards* are the mechanisms most likely to achieve such impact, although this metric was not a significant factor in the selection of these themes. The successful *Save Snapshot Serengeti* campaign achieved a significant impact, reaching thousands of users a day and even the general public, by setting goals for users to share project materials [21]. While goals and rewards are common in online CS projects, we notice a focus on microtask completion, rather than community engagement goals and rewards. We suggest that increased use of goal setting, particularly community-based goals, as well as increased use of community-based rewards will lead to increased project engagement. In particular, meta-campaigns and social media sharing campaigns are a simple, yet effective method for reaching people outside of the community of participants [59].

Engagement (ii) – Number of contributions received

A further measure of successful engagement is the number of classifications recieved by a project. Any CS project must receive a minimum number of classifications to meet its goal and facilitate scientific research. Moreover, increasing the volume of contributions received by increasing engagement may confer further advantages, such as allowing for increased accuracy through the aggregation of results (see for example: [53]). Similarly, contributions to discussion platforms can serve as a second path to scientific discoveries (see for example: [5]). Our results show that the task visibility mechanisms used support microtask completion, but are less effective for community contributions. This is reflected in the large percentage (>90%) of *Zooniverse Talk* comments which lack responses [29]. It is our view that increased use of community-related task visibility mechanisms and ensuring equal rewards for both microtask and community contribution will increase the number of contributions received within community features, increasing volunteer engagement. Furthermore, effective use of feedback can positively contribute to the number of contributions received from volunteers in VCS projects [1, 51].

Accuracy and Quality – Validity of results gained

In order for CS results to be used for scientific research, they must be accurate. CS projects must therefore positively reinforce correct contributions from volunteers. Within our survey, both *feedback* and *rewards* were identified as playing the greatest role in allowing project design teams to ensure the accuracy of contributions, not only ensuring that results are corrected in real time, but also correcting participants to ensure the accuracy of future contributions. Feedback and reward mechanisms are popular components in the CS projects surveyed, with a general focus on task-contingent, quantity-based feedback. Performance-contingent feedback, however, was less common. Instead of focusing solely on quantity of microtasks completed, we recommend shifting these mechanisms to assess quality of contributions, as a means of improving project accuracy. We note that projects must often assess accuracy before publishing data, by comparison with gold standards or through calculation of other metrics (see for example: [53, 27]). A simple mechanism for providing performance-based feedback would be to carry out such a process earlier, in tandem with volunteer contributions, allowing for the results to be shared with volunteers on an ongoing basis, in the form of performance-based feedback. While generating such feedback can be complex, we note that even relatively simple feedback such as points can be a useful measure for volunteers when determining how well they are performing and when attempting to improve [11]. Wider community feedback can also be an important source of increased accuracy. The project iSpotNature, which relies exclusively on community-derived feedback, identified increases in the accuracy of metadata attached to submissions in 57% of cases [48]. Conversely, such an approach alone may be insufficient to ensure accuracy - despite the relatively high overall accuracy (96.6%) achieved by the Snapshot Serengeti project, which also features only community-based feedback, certain species feature much lower accuracy rates, in cases as low as 33%, with rarer species more likely to result in false positives [53]. We observe that determining the effectiveness of such feedback is further complicated by the difficulties in receiving responses identified within project literature (see for example: [29, 59]).

Design Recommendations

Designing online citizen science projects is a complex process, with individual design decisions impacting a number of factors, including volunteer engagement and motivations, data quantity and quality and the research outcomes of a project.

In terms of task visibility, we identified 6 mechanisms among the surveyed projects for identifying discussions in need of contributions. However, in contrast to task-related mechanisms, these mechanisms did not ensure equal attention is given to each task, as these mechanisms provided little support for volunteers in identifying those threads most in need of attention. The most common mechanism, *notification of most recent activity* is unsuited for the asynchronous nature of many of the platforms identified, requiring volunteers to observe the platform around the time a post is made to be able to see it. Similarly, while *free selection of discussion threads* allows volunteers to find discussions of interest to themselves, the large number of posts involved in many projects increases the likelihood that some posts will go unseen. In just 7 months, Snapshot Serengeti generated 39,250 discussion posts, while SpaceWarps generated 20,978 posts in 2 months [29]. Other mechanisms require users to specifically seek out threads and subscribe in order to see further responses.

As an alternative, we propose enabling volunteers to order posts according to the number of replies that a post has received. Evidence from both Zooniverse and FoldIt suggests that volunteers fulfil specific roles when engaging in community discussions, including answering and contributing to questions and open discussions [9, 58]. By simplifying the process of finding such discussions, we believe that volunteers will be able to reduce thread response-times and the number of unanswered threads with minimal input from project scientists.

Setting suitable goals for challenges is a difficult process. If goals are too simple or deemed unachievable by volunteers, then they can negatively impact the number of classifications volunteers submit [25, 28]. Levels of engagement can be extremely unpredictable, as in the case of the Andromeda Project, a Zooniverse citizen science project where volunteers successfully completed over a million classifications in just two weeks, a feat that was expected to take two months [17].

We therefore propose the use of meta-challenges, surveys and community-based competitions, instead of challenges explicitly linked to task-completion. These challenges have shown to be effective in attracting volunteers to projects and increasing the completion of task classifications. Furthermore, such challenges can provide fund raising opportunities, gathering resources for projects while at the same time not requiring the heavier time or money investments that may be associated with classification challenges. This is particularly valuable given that task-based challenges were often coupled with physical prizes and that successful use of such challenges requires time investment and careful monitoring from community moderators and design teams [22].

Differing forms of feedback serve different, but equally important roles, in positively re-enforcing volunteer behaviour. Performance-contingent feedback is essential for complex tasks, allowing volunteers to identify whether they are contributing correctly [11]. Task-contingent feedback is equally valuable, in reassuring volunteers that their contributions are valued and used by scientists [51]. Furthermore, both forms of feedback serve to reinforce the function of goals, allowing volunteers to follow their progress with respect to other volunteers and goal deadlines [22]. Feedback is particularly valuable in VCS, where doubt has been cast on the use of tutorials as a means of training volunteers. Starr et al note that video-based, online tutorials can be as effective as in person training for citizen science tasks [52]. However, Newman et al note that such training is unsuitable for more complex skills and tools, which volunteers struggle with even after completing the tutorial process [34]. This is further exacerbated by the unwillingness of many volunteers to complete the tutorial process, reducing the size of the community [10], or necessitating the use of non-compulsory tutorials [59].

We recommend that projects deliver both task- and performance-contingent feedback. Task-contingent feedback should be delivered predominantly through automated calculations, providing volunteers with dashboard-style statistics on project completion, or with more competitive projects, through leaderboards and point calculations. This allows volunteers to receive updates in real-time, while also reducing the overall workload for project scientists. Performance-contingent feedback is more complex, as noted, due to the lack of 'correct' responses. We suggest a trade-off between the accuracy of feedback and the level of workload required of project scientists. In its simplest, but least accurate form, projects can compare responses with the majority opinion. More accurate feedback can be generated by creating a gold-standard set of images with which volunteers responses can be compared, but such an approach requires the investment of time before projects launch. Furthermore, as the number of entities within a project increases, further attention is required from project scientists if the level of feedback offered is to remain consistent.

As with goals, rewards have the potential to negatively impact volunteer behaviour, reducing motivation and encouraging users to game the system to receive maximum rewards from minimum effort [10, 22]. Such effects are associated with specific forms of reward: physical rewards and task-contingent reward structures are more likely to encourage such behaviour than status rewards, particularly for those who are less invested in the community [22].

We therefore propose that rewards should be performance-contingent, encouraging volunteers to create quality submissions, rather than a large quantity of lower quality submissions. In this way, rewards can serve as a further feedback mechanism, re-enforcing positive behaviour and allowing volunteers to monitor the quality of their own contributions. Status rewards should be utilised to reduce the likelihood of negative behaviours, while also reducing the resources required to produce rewards - physical rewards are likely to be costly, which is problematic for VCS projects. By monitoring the status awarded to volunteers, project scientists can identify those volunteers who are most dedicated to the project and confer on them specific roles such as *moderator*. Evidence from Zooniverse's Talk system suggests that volunteer moderators can be highly effective in identifying and flagging topics which require attention from science teams, reducing the effort required of project scientists while helping to ensure that discussions do not go unanswered [59].

Limitations

The facets discussed within this work are only a small subset of the vast number of dimensions for online community success discussed within the literature. While we have selected those deemed most salient, it is clear that many other factors must be considered in designing and building successful community-based online citizen science projects. One such area is the implementation of gamification. A number of gamified aspects have been identified within this survey, such as user motivation, leaderboards, point scores, badges, and ranks. This is particularly significant within the Games With A Purpose such as *EteRNA*, *Phylo* and *FoldIt*. While our work has focused on the use of online community mechanisms, other studies suggest gamification may also play a key role. Mekler has demonstrated that gamification elements affect the level and form which engagement in CS takes, while Bowser et al suggest gamification may be key to attracting demographics such as millenials to CS projects [31, 4].

We also are aware of the limitations pertaining to identifying impacts on project success. Due to the observational methodology utilised within this work, it is not possible to directly quantify the effect that the use or lack of a given mechanism has had on particular success metrics. However, identifying quantitative measures for certain design decisions is also a difficult process even were a differing methodology to be employed – particularly in citizen science, where a number of compounding factors such as volunteer interests may exist. We believe this is an area for further research, although such research will need to consider precise measures for the effects of such decisions.

RELATED WORK

In this section we discuss related work which has contributed to our research. We highlight the contributions that these studies have made to the research process and outlining the ways in which our work builds on and otherwise deviates from the existing literature.

A similar study concerning factors impacting the quality and quantity of contributions to online citizen science projects was conducted by Nov et al [35]. The authors looked at 3 systems, *Stardust@home, The Citizen Weather Observer Program (CWOP) and The Berkeley Open Infrastructure for Network Computing (BOINC)*. In particular, the analysis conducted by Nov et al concentrated on the impact of individual motivational factors and forms of motivation on levels of contribution and the quality of contribution, in terms of Collective, Norm-Oriented and Intrinsic Motives, as well as Reputation. While all four motives were found affect the quantity of contributions received, only collective motives and reputation were found to positively influence the quality of submissions. We note the similar questions raised by this research and utilised this study in the literature review which formed the evidence basis for the analysis outlined within the discussion section of this paper. However, our research differs also differs greatly from that of Nov et al, drawing on a larger selection and wider range of projects. Furthermore, our work has an online communities and design focus, informed by the design decisions underlying the projects studied and the wider online community and citizen science research informing and describing the results of such decision.

Kullenberg and Kasperowski conducted a large-scale analysis of citizen science literature, drawn from the Web of Science database [23]. This analysis drew on two datasets of publications, one comprised of 1935 items and one of 633 items, in order to conceptualise citizen science and the processes and aims associated with it. This work provided important insights into the literature review process used, including the keywords selected and identifying methods for removing irrelevant papers, false positives and negatives and other outliers generated through the search methods used. While our work shares some similarities with that of Kullenberg and Kasperowski, we drew on a comparatively smaller, but more varied body of literature, using a larger range of databases. In addition, our search terms focused exclusively on online citizen science projects, in contrast with the more general focus of the search conducted by Kullenberg and Kasperowski.

CONCLUSIONS

In this paper we performed a systematic structured walkthrough of 48 citizen science projects to investigate common features implemented in such platforms. Based on our analysis, we identified a number of relevant design claims for motivating user contributions, across the themes of task visibility, goals, feedback, and rewards.

Online citizen science projects serve as a unique form of online community and an understanding of such systems continues to emerge. As with all online communities, citizen science projects face challenges with regard to encouraging contributions from users, both in the form of the microtask component of a project, and community participation. Citizen science communities face further unique challenges with regard to ensuring the validity of data and justifying the use of a crowdsourced, citizen science-based approach.

Our analysis has demonstrated a close connection between task and community aspects of CS projects, previously considered to be separate dimensions. We have further identified links between the use of online community design principles and CS project success metrics, although we recommend that further consideration should be given to how design decisions and the inclusion of features may impact these metrics. One key area for future work will be exploring quantitative measures for specific design decisions, to allow for more informed decision making in CS design.

ACKNOWLEDGEMENTS

This work was supported by the Web Science Centre for Doctoral Training at the University of Southampton, funded by the UK Engineering and Physical Sciences Research Council (EPSRC) under grant number EP/G036926/1; by the research project *SOCIAM: The Theory and Practise of Social Machines* funded by the EPSRC under grant number EP/J017728/2 and by the research project *STARS4ALL* funded by the European Commission under grant number 688135.

REFERENCES

1. Maria Aristeidou, Eileen Scanlon, and Mike Sharples. 2015. Weather-it: evolution of an online community for citizen inquiry. In *Proceedings of the 15th International Conference on Knowledge Technologies and Data-driven Business*. ACM, 13.
2. Verb Corner Blog. Retrieved Sep 2015. VerbCorner: New and improved with surprise bonuses - VerbCorner blog. `http://gameswithwords.fieldofscience.com/2013/07/verbcorner-new-and-improved-with.html`. (Retrieved Sep 2015).
3. Rick Bonney, Caren B Cooper, Janis Dickinson, Steve Kelling, Tina Phillips, Kenneth V Rosenberg, and Jennifer Shirk. 2009. Citizen science: a developing tool for expanding science knowledge and scientific literacy. *BioScience* 59, 11 (2009), 977–984.
4. Anne Bowser, Derek Hansen, Yurong He, Carol Boston, Matthew Reid, Logan Gunnell, and Jennifer Preece. 2013. Using gamification to inspire new citizen science volunteers. In *Proceedings of the First International Conference on Gameful Design, Research, and Applications*. ACM, 18–25.
5. Carolin Cardamone, Kevin Schawinski, Marc Sarzi, Steven P Bamford, Nicola Bennert, CM Urry, Chris Lintott, William C Keel, John Parejko, Robert C Nichol, and others. 2009. Galaxy Zoo Green Peas: discovery of a class of compact extremely star-forming galaxies. *Monthly Notices of the Royal Astronomical Society* 399, 3 (2009), 1191–1205.
6. Denis Couvet, Frédéric Jiguet, Romain Julliard, Harold Levrel, and Anne Teyssèdre. 2008. Enhancing citizen contributions to biodiversity science and public policy. *Interdisciplinary science reviews* 33, 1 (2008), 95–103.
7. Joe Cox, Eun Young Oh, Brooke Simmons, Chris Lintott, Karen Masters, Anita Greenhill, Gary Graham, and Kate Holmes. 2015. Defining and Measuring Success in Online Citizen Science: A Case Study of Zooniverse Projects. *Computing in Science & Engineering* 17, 4 (2015), 28–41.
8. Kevin Crowston and Nathan R Prestopnik. 2013. Motivation and data quality in a citizen science game: A design science evaluation. In *System Sciences (HICSS), 2013 46th Hawaii International Conference on*. IEEE, 450–459.
9. Vickie Curtis. 2014. Online citizen science games: opportunities for the biological sciences. *Applied & Translational Genomics* 3, 4 (2014), 90–94.
10. Vickie Curtis. 2015a. Motivation to Participate in an Online Citizen Science Game A Study of Foldit. *Science Communication* 37, 6 (2015), 723–746.
11. Vickie Curtis. 2015b. *Online citizen science projects: an exploration of motivation, contribution and participation*. Ph.D. Dissertation. The Open University.
12. Stuart Dunn and Mark Hedges. 2013. Crowd-sourcing as a component of humanities research infrastructures. *International Journal of Humanities and Arts Computing* 7, 1-2 (2013), 147–169.
13. Muki Haklay. 2010. Geographical citizen science–Clash of cultures and new opportunities. (2010).
14. Jeff Howe. 2006. The rise of crowdsourcing. *Wired magazine* 14, 6 (2006), 1–4.
15. Alicia Iriberri and Gondy Leroy. 2009. A life-cycle perspective on online community success. *ACM Computing Surveys (CSUR)* 41, 2 (2009), 11.
16. Nicholas R Jennings, Luc Moreau, D Nicholson, S Ramchurn, S Roberts, T Rodden, and Alex Rogers. 2014. Human-agent collectives. *Commun. ACM* 57, 12 (2014), 80–88.
17. Cliff Johnson. 2012. Round #1 Complete! `https://blog.andromedaproject.org/2012/12/21/round-1-complete/`. (2012).
18. Charles R Kerton, Grace Wolf-Chase, Kim Arvidsson, Chris J Lintott, and Rob J Simpson. 2015. The Milky Way Project: What Are Yellowballs? *The Astrophysical Journal* 799, 2 (2015), 153.
19. Laure Kloetzer, Daniel Schneider, Charlene Jennett, Ioanna Iacovides, Alexandra Eveleigh, Anna Cox, and Margaret Gold. 2014. Learning by volunteer computing, thinking and gaming: What and how are volunteers learning by participating in Virtual Citizen Science? *Changing Configurations of Adult Education in Transitional Times* (2014), 73.
20. Margaret Kosmala. 2012. We Need an I Dont Know Button! `http://blog.snapshotserengeti.org/2012/12/14/we-need-an-i-dont-know-button/`. (2012).
21. Margaret Kosmala. Retrieved Sep 2015. Analysis of 'Save Snapshot Serengeti' - Snapshot Serengeti blog. `http://blog.snapshotserengeti.org/2013/08/28/analysis-of-save-snapshot-serengeti/`. (Retrieved Sep 2015).

22. Robert E Kraut, Paul Resnick, Sara Kiesler, Moira Burke, Yan Chen, Niki Kittur, Joseph Konstan, Yuqing Ren, and John Riedl. 2012. *Building successful online communities: Evidence-based social design*. Mit Press.

23. Christopher Kullenberg and Dick Kasperowski. 2016. What Is Citizen Science?–A Scientometric Meta-Analysis. *PloS one* 11, 1 (2016), e0147152.

24. Cliff Lampe and Erik Johnston. 2005. Follow the (slash) dot: effects of feedback on new members in an online community. In *Proceedings of the 2005 international ACM SIGGROUP conference on Supporting group work*. ACM, 11–20.

25. Kimberly Ling, Gerard Beenan, Pamela Ludford, Xiaoqing Wang, Klarissa Chang, Xin Li, Dan Cosley, Dan Frankowski, Loren Terveen, Al Mamunur Rashid, and others. 2005. Using social psychology to motivate contributions to online communities. *Journal of Computer-Mediated Communication* 10, 4 (2005).

26. Chris Lintott. Retrieved Sep 2015. Optimising For Interest: Why People Aren't Machines. `https://blog.zooniverse.org/2013/03/29/optimizing-for-interest-why-people-aren't-machines/`. (Retrieved Sep 2015).

27. Chris Lintott, Kevin Schawinski, Steven Bamford, Anže Slosar, Kate Land, Daniel Thomas, Edd Edmondson, Karen Masters, Robert C Nichol, M Jordan Raddick, and others. 2011. Galaxy Zoo 1: data release of morphological classifications for nearly 900 000 galaxies. *Monthly Notices of the Royal Astronomical Society* 410, 1 (2011), 166–178.

28. Edwin A Locke, Gary P Latham, and Miriam Erez. 1988. The determinants of goal commitment. *Academy of management review* 13, 1 (1988), 23–39.

29. Markus Luczak-Roesch, Ramine Tinati, Elena Simperl, Max Van Kleek, Nigel Shadbolt, and Robert Simpson. 2014. Why won't aliens talk to us? Content and community dynamics in online citizen science. (2014).

30. Kurt Luther, Scott Counts, Kristin B Stecher, Aaron Hoff, and Paul Johns. 2009. Pathfinder: an online collaboration environment for citizen scientists. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 239–248.

31. Elisa D Mekler, Florian Brühlmann, Klaus Opwis, and Tuch Alexandre N. 2013. Disassembling gamification: the effects of points and meaning on user motivation and performance. In *CHI'13 extended abstracts on human factors in computing systems*. ACM, 1137–1142.

32. Pietro Michelucci. 2013. *Handbook of human computation*. Springer.

33. Gabriel Mugar, Carsten Østerlund, Katie DeVries Hassman, Kevin Crowston, and Corey Brian Jackson. 2014. Planet hunters and seafloor explorers: legitimate peripheral participation through practice proxies in online citizen science. In *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing*. ACM, 109–119.

34. Greg Newman, Alycia Crall, Melinda Laituri, Jim Graham, Tom Stohlgren, John C Moore, Kris Kodrich, and Kirstin A Holfelder. 2010. Teaching citizen science skills online: Implications for invasive species training programs. *Applied Environmental Education and Communication* 9, 4 (2010), 276–286.

35. Oded Nov, Ofer Arazy, and David Anderson. 2014. Scientists@Home: what drives the quantity and quality of online citizen science participation? *PloS one* 9, 4 (2014), e90375.

36. Johan Oomen and Lora Aroyo. 2011. Crowdsourcing in the cultural heritage domain: opportunities and challenges. In *Proceedings of the 5th International Conference on Communities and Technologies*. ACM, 138–149.

37. OpenScientist. Retrieved Sep 2015. Finalizing a Definition of Citizen Science and Citizen Scientists. `http://www.openscientist.org/2011/09/finalizing-definition-of-citizen.html`. (Retrieved Sep 2015).

38. Jenny Preece. 2001. Sociability and usability in online communities: Determining and measuring success. *Behaviour & Information Technology* 20, 5 (2001), 347–356.

39. Nathan R Prestopnik. 2012. Citizen science case study: Galaxy Zoo/Zooniverse. (2012).

40. C Aaron Price and Hee-Sun Lee. 2013. Changes in participants' scientific attitudes and epistemological beliefs during an astronomical citizen science project. *Journal of Research in Science Teaching* 50, 7 (2013), 773–801.

41. M Jordan Raddick, Georgia Bracey, Karen Carney, Geza Gyuk, Kirk Borne, John Wallin, Suzanne Jacoby, and Adler Planetarium. 2009. Citizen science: status and research directions for the coming decade. *AGB Stars and Related Phenomenastro 2010: The Astronomy and Astrophysics Decadal Survey* (2009), 46P.

42. Yuqing Ren, Robert Kraut, and Sara Kiesler. 2007. Applying common identity and bond theory to design of online communities. *Organization studies* 28, 3 (2007), 377–408.

43. Amy Robinson. Retrieved Sep 2015. How are Points Calculated in EyeWire? - EyeWire blog. `http://blog.eyewire.org/how-are-points-calculated-in-eyewire`. (Retrieved Sep 2015).

44. Helen E Roy, Michael JO Pocock, Chris D Preston, David B Roy, J Savage, JC Tweddle, and LD Robinson. 2012. Understanding citizen science and environmental monitoring: final report on behalf of UK Environmental Observation Framework. (2012).

45. Meg Schwamb. Retrieved Apr 2016. 2 Years On from BBC Stargazing Live - Planet Four blog. `http://blog.planetfour.org/2015/03/18/2-years-on-from-bbc-stargazing-live/`. (Retrieved Apr 2016).

46. Megan E Schwamb, Jerome A Orosz, Joshua A Carter, William F Welsh, Debra A Fischer, Guillermo Torres, Andrew W Howard, Justin R Crepp, William C Keel,

Chris J Lintott, and others. 2013. Planet hunters: A transiting circumbinary planet in a quadruple star system. *The Astrophysical Journal* 768, 2 (2013), 127.

47. Jonathan Silvertown. 2009. A new dawn for citizen science. *Trends in ecology & evolution* 24, 9 (2009), 467–471.

48. Jonathan Silvertown, Martin Harvey, Richard Greenwood, Mike Dodd, Jon Rosewell, Tony Rebelo, Janice Ansine, and Kevin McConway. 2015. Crowdsourcing the identification of organisms: A case-study of iSpot. *ZooKeys* 480 (2015), 125.

49. Rob Simpson. Retrieved Apr 2016. Stargazing Live: The Results Are In. https://blog.zooniverse.org/2014/01/09/stargazing-live-the-results-are-in/. (Retrieved Apr 2016).

50. Arfon Smith. Retrieved Sep 2015. How the Zooniverse Works: Keeping it Personal. https://blog.zooniverse.org/2013/07/23/keeping-it-personal/. (Retrieved Sep 2015).

51. James Sprinks, Robert Houghton, Steven Bamford, and Jeremy Morley. 2015. Citizen Scientists: The Importance of Being Needed and not Wasted. In *Proceedings of the 2015 Annual Symposium on Computer-Human Interaction in Play*. ACM.

52. Jared Starr, Charles M Schweik, Nathan Bush, Lena Fletcher, Jack Finn, Jennifer Fish, and Charles T Bargeron. 2014. Lights, camera... citizen science: assessing the effectiveness of smartphone-based video training in invasive species plant identification. *PloS one* 9, 11 (2014), e111433.

53. Alexandra Swanson, Margartet Kosmala, Chris Lintott, Robert Simpson, Arfon Smith, and Craig Packer. 2015. Snapshot Serengeti, high-frequency annotated camera trap images of 40 mammalian species in an African savanna. *Scientific data* 2 (2015).

54. EteRNA Team. Retrieved Sep 2015a. EteRNA wiki - EteRNA. http://eternawiki.org/wiki/index.php5/EteRNA. (Retrieved Sep 2015).

55. Stardust@home Team. Retrieved Sep 2015b. Interstellar Candidates - Stardust@home. http://stardustathome.ssl.berkeley.edu/news/candidates/. (Retrieved Sep 2015).

56. Ellinore J Theobald, Ailene K Ettinger, Hillary K Burgess, Lauren B DeBey, Natalie R Schmidt, Halley E Froehlich, Christian Wagner, Janneke HilleRisLambers, Joshua Tewksbury, Melanie Harsch, and Julia K Parrish. 2015. Global change and local solutions: Tapping the unrealized potential of citizen science for biodiversity research. *Biological Conservation* 181 (2015), 236–244.

57. Ramine Tinati, Markus Luczak-Rösch, Elena Simperl, and Wendy Hall. 2016. "Science is Awesome":Studying Participation in a Citizen Science Game. *ACM Web Science Conference 2016* (2016).

58. Ramine Tinati, Markus Luczak-Rösch, Elena Simperl, Wendy Hall, and Nigel Shadbolt. 2015a. '/Command' and conquer: analysing discussion in a citizen science game. *ACM Web Science Conference 2015* (2015).

59. Ramine Tinati, Max Van Kleek, Elena Simperl, Markus Luczak-Rösch, Robert Simpson, and Nigel Shadbolt. 2015b. Designing for Citizen Data Analysis: A Cross-Sectional Case Study of a Multi-Domain Citizen Science Platform. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. ACM, 4069–4078.

60. Tamsyn P Waterhouse. 2013. Pay by the bit: an information-theoretic metric for collective human judgment. In *Proceedings of the 2013 conference on Computer supported cooperative work*. ACM, 623–638.

61. Andrea Wiggins and Kevin Crowston. 2010. Developing a conceptual model of virtual organisations for citizen science. *International Journal of Organisational Design and Engineering* 1, 1-2 (2010), 148–162.

62. Andrea Wiggins and Kevin Crowston. 2011. From conservation to crowdsourcing: A typology of citizen science. In *System Sciences (HICSS), 2011 44th Hawaii international conference on*. IEEE, 1–10.

63. Andrea Wiggins and Kevin Crowston. 2012. Goals and tasks: Two typologies of citizen science projects. In *System Science (HICSS), 2012 45th Hawaii International Conference on*. IEEE, 3426–3435.

64. Haiyi Zhu, Robert Kraut, and Aniket Kittur. 2012. Organizing without formal organization: group identification, goal setting and social modeling in directing online production. In *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work*. ACM, 935–944.

APPENDIX

List of Projects Surveyed

Project Name	Project URL	Community Features
Annotate	https://anno.tate.org.uk	Talk (Zooniverse Custom Platform)
Asteroid Mappers	https://cosmoquest.org/?application=vesta_mappers/	Discussion Board Forum
Asteroid Zoo	http://www.asteroidzoo.org	Talk (Zooniverse Custom Platform)
Bug Guide	http://bugguide.net/	Discussion Board Forum
Chicago Wildlife Watch	http://www.chicagowildlifewatch.org	Talk (Zooniverse Custom Platform)
Chimp and See	http://www.chimpandsee.org	Talk (Zooniverse Custom Platform)
Condor Watch	http://www.condorwatch.org	Talk (Zooniverse Custom Platform)
Cyclone Centre	http://www.cyclonecenter.org	Talk (Zooniverse Custom Platform)
Disk Detective	http://www.diskdetective.org	Talk (Zooniverse Custom Platform)
EteRNA	http://www.eternagame.org	Live Instant Messenger Chat, Discussion Board Forum, Wiki
EyeWire	http://eyewire.org/	Live Instant Messenger Chat, Discussion Board Forum, Wiki
Floating Forests	http://www.floatingforests.org	Talk (Zooniverse Custom Platform)
FoldIt	http://fold.it/portal/	Discussion Board Forum, Wiki
Fossil Finder	http://www.zooniverse.org/projects/adrianevans/fossil-finder/	Talk (Zooniverse Custom Platform)
Galaxy Zoo	http://www.galaxyzoo.org	Talk (Zooniverse Custom Platform)
Galaxy Zoo Bar Lengths	http://www.zooniverse.org/projects/vrooje/galaxy-zoo-bar-lengths/	Talk (Zooniverse Custom Platform)
Herbaria@Home	http://herbariaunited.org/atHome/	Discussion Board Forum
Higgs Hunters	http://www.higgshunters.org	Talk (Zooniverse Custom Platform)
Instant Wild	http://www.edgeofexistence.org/instantwild/	Comment Listing
iSpotNature	http://www.ispotnature.org	Discussion Board Forum
Landscape Watch Hampshire	http://www.hampshire.landscapewatch.com/	Discussion Board Forum
Mars Mappers	https://cosmoquest.org/?application=mars_simply_craters	Discussion Board Forum
Mercury Mappers	https://cosmoquest.org/projects/mercury_mappers	Discussion Board Forum
Milky Way Project	http://www.milkywayproject.org/	Talk (Zooniverse Custom Platform)
Moon Mappers	https://cosmoquest.org/?application=simply_craters	Discussion Board Forum
Notes From Nature	http://www.notesfromnature.org	Talk (Zooniverse Custom Platform)
Old Weather	http://www.oldweather.org	Discussion Board Forum
Operation War Diary	http://www.operationwardiary.org	Talk (Zooniverse Custom Platform)
Orchid Observers	http://www.orchidobservers.org	Talk (Zooniverse Custom Platform)
Penguin Watch	http://www.penguinwatch.org	Talk (Zooniverse Custom Platform)
Phylo	http://phylo.cs.mcgill.ca/	Discussion Board Forum
Planet Four	http://www.planetfour.org	Talk (Zooniverse Custom Platform)
Planet Four: Terrains	http://www.zooniverse.org/projects/mschwamb/planet-four-terrains/	Talk (Zooniverse Custom Platform)
Planet Hunters	http://www.planethunters.org	Talk (Zooniverse Custom Platform)
Plankton Portal	http://planktonportal.org/	Talk (Zooniverse Custom Platform)
Radio Galaxy Zoo	http://radio.galaxyzoo.org	Talk (Zooniverse Custom Platform)
Science Gossip	http://www.sciencegossip.org	Talk (Zooniverse Custom Platform)
Season Spotter Image Marking	http://www.zooniverse.org/projects/kosmala/season-spotter-image-marking	Talk (Zooniverse Custom Platform)
Season Spotter Questions	http://www.zooniverse.org/projects/kosmala/season-spotter-questions	Talk (Zooniverse Custom Platform)
Snapshot Serengeti	http://snapshotserengeti.org/	Talk (Zooniverse Custom Platform)
SpaceWarps	http://spacewarps.org/	Talk (Zooniverse Custom Platform)
Stardust@Home	http://stardustathome.ssl.berkeley.edu/	Discussion Board Forum
Sunspotter	http://www.sunspotter.org	Talk (Zooniverse Custom Platform)
Verb Corner	http://gameswithwords.org/VerbCorner	Discussion Board Forum
Whales As Individuals	http://www.zooniverse.org/projects/tedcheese/whales-as-individuals	Talk (Zooniverse Custom Platform)
Wildcam Gorongosa	http://www.wildcamgorongosa.org	Talk (Zooniverse Custom Platform)
Wildebeest Watch	http://www.zooniverse.org/projects/aliburchard/wildebeest-watch/	Talk (Zooniverse Custom Platform)
Worm Watch Lab	http://www.wormwatchlab.org	Talk (Zooniverse Custom Platform)

“A Game Without Competition is Hardly a Game”: The Impact of Competitions on Player Activity in a Human Computation Game

Neal Reeves and **Peter West** and **Elena Simperl**
Web and Internet Science, Electronics and Computer Science, University of Southampton

Abstract

Virtual citizen science (VCS) projects enable new forms of scientific research using crowdsourcing and human computation to gather and analyse large-scale datasets. To attract and sustain the number of participants and levels of participation necessary to achieve research aims, some VCS projects have introduced game elements such as competitions to tasks. However, we still know very little about how some game elements, particularly competitions, influence participation rates. To investigate the impact of game elements on player engagement, we conducted a two-part mixed-methods study of EyeWire, a VCS game. First, we interviewed EyeWire designers to understand their rationale for introducing competitions. Guided by their answers, we analysed two datasets of EyeWire user task contributions and chat logs to assess the effectiveness of competitions in achieving designers' goals. Our findings contribute to the growing understanding of how competitions influence participant activity in human computation initiatives and socio-technical systems such as VCS.

Introduction

Virtual citizen science (VCS) projects are a form of human computation, which draws on large-scale crowdsourcing methodologies to gather and analyse datasets for scientific research purposes (Wiggins and Crowston 2011). As modern technology has enabled and demanded the collection of large volumes of data and images, traditional scientific workflows where small teams of scientists collect and analyse data have becoming increasingly insufficient (Lintott et al. 2008). The use of citizen science methodologies has been increasingly suggested as a solution to such issues, with participants preparing datasets or training algorithms for future research use (Khatib et al. 2011; Lintott et al. 2008). Relying on the 'wisdom of crowds', VCS initiatives serve as socio-technical systems, which draw on a combination of human and machine agents to generate and process data (Jennings et al. 2014; Michelucci 2013). Furthermore, projects serve as online communities, where participants may use computer-mediated communication platforms to interact with one another and further engage with scientific research (Luczak-Roesch et al. 2014; Østerlund et al. 2014; Oliveira, Jun, and Reinecke 2017).

However, as with other human computation initiatives, designing suitable VCS projects remains a pervasive research topic. Projects must balance a number of trade-offs to generate sufficiently accurate data, while attracting and supporting large communities of participants in undertaking unfamiliar and potentially complex tasks (Tinati et al. 2015b). Datasets may consist of hundreds of thousands to millions of artefacts, each potentially requiring analysis by multiple participants to ensure the validity of responses (Tinati et al. 2015b; Wiggins et al. 2011). As with many other web-based collaborative communities, a minority of participants often produce the majority of contributions, with a large proportion of users contributing for single, brief sessions (Ponciano et al. 2014; Sauermann and Franzoni 2015). To address problems, an increasing number of projects are turning to gamification and game elements as a means to attract and sustain volunteer engagement with VCS and increase levels of participation (Bowser et al. 2013; Curtis 2014; Reeves et al. 2017).

Yet the use of gamification for human computation is not without its challenges. The effectiveness of game elements depends on the context of platforms and the personality characteristics of participants (Jia et al. 2016; 2017). Gamified elements often introduce extrinsic reward mechanisms such as points, badges and physical prizes (Reeves et al. 2017; Seaborn and Fels 2015). However, studies have shown that participants have different motivations; while some align well with these elements (Jia et al. 2016), other volunteers' intrinsic motivations are negatively impacted by the use of extrinsic factors (Greenhill et al. 2016). This is problematic in VCS, where projects predominantly rely on volunteers' intrinsic motivations to encourage and motivate contributions (Cox et al. 2015; Rotman et al. 2012). In particular, we know less about the effects of competitive game elements such as global leaderboards and contests. Introducing competitive elements has been observed to make tasks stressful and demotivating, as individual players gain unassailable leads over fellow participants (Eveleigh et al. 2013). The impact of point-based game design and the use of interplayer competitions has been suggested to be minimal, in contrast to narrative and framing devices (Lieberoth 2015; Mekler et al. 2013a; Prestopnik and Tang 2015). Specific to VCS, game elements risk presenting simplistic ideas of scientific research (Sandbrook, Adams, and Monteferri 2015), negatively influencing engagement in VCS (Eveleigh et al. 2013; Ponti, Hillman, and Stankovic 2015).

In this paper, we explore the impact of competition events in *EyeWire*[1], a Virtual Citizen Science Game With a Purpose. We first conduct a series of structured reflection interviews with six key members of the EyeWire project team, in order to identify the factors motivating the implementation of competition events within the project. Using the results of this interview process as an analytical framework, we then assess the effectiveness of competitions held within a one year period compared with non-competitive periods, across factors such as number of contributions, number of participants, performance of additional tasks and social interaction. Our work builds on previous research in this area by drawing on the experiences of VCS designers and offering quantitative evidence of the impacts of competitions on player activity levels in VCS.

Background and Related Work

Citizen Science

Citizen science (CS) describes the engagement of volunteer participants in the scientific research process (Dickinson, Zuckerberg, and Bonter 2010). Volunteers undertake several activities, including data collection, classification and analysis (Tinati et al. 2015b). In some forms of citizen science, participants may even co-design experiments and research questions, drawing conclusions and publishing findings (Haklay 2013). As access to the Web has grown, so has a new form of citizen science: *Virtual Citizen Science*. In this form, rather than gathering data and making records, participants contribute their knowledge through crowd-sourced human computation activities, delivered through web-based and mobile portals (Wiggins and Crowston 2011).

As a human computation process, VCS is characterised by a number of tradeoffs between factors such as engagement, ease of use and data quality (Aoki et al. 2017; Sprinks et al. 2015b). For example, while VCS introduces new concepts and tasks, completion rates of tutorials are low, resulting in trade-offs or the need to support human computation through algorithms and software (Mattos et al. 2014). Findings from a survey of participant motivations in the Zooniverse platform carried out by Sprinks et al suggest that players require feedback to resolve task performance concerns and to feel valued (Sprinks et al. 2015a).

VCS is also highly dependent on the suitability of tasks provided to volunteers. Breaking overarching macrotasks into smaller, repeateable microtask actions is a distinct challenge - one that has been shown to increase the overall time taken to complete tasks, yet lead to an increase in the quality of outputs and user experience (Cheng et al. 2015). Design decisions must also consider a wide variety of factors and draw on a number of skillsets and array of disciplinary knowledge. Tinati et al. synthesised design recommendations from interview sessions with key members of the design team responsible for the Zooniverse platform (Tinati et al. 2015b). These recommendations show the interplay of complex and potentially unpredictable factors in motivating VCS engagement. Beyond engagement, VCS projects are also subject to a variety of data quality and accuracy concerns. Participants can nonetheless be a valuable resource for ensuring the validity of responses, but this increases the workload expected of the community and thus requires further engagement. For example, repetition of individual microtasks by multiple users is a common form of data validation in VCS (Wiggins et al. 2011).

Gamification

Gamification describes the use of "game elements in non-gaming contexts" (Deterding et al. 2011). We make the distinction between gamification and the similar, yet distinct, *Games With A Purpose*, where participants carry out human computation tasks as part of a casual game experience. (Siu, Zook, and Riedl 2014). While human computation initiatives such as Citizen Science projects may make use of gamified elements (such as points and rewards) to encourage player engagement with tasks, it is the diegetic nature of the elements in Games With A Purpose which sets them apart from other project varieties (Prestopnik, Crowston, and Wang 2014). Gamification has been employed in a wide array of contexts, from education and research to business and marketing (Seaborn and Fels 2015) and a significant focus of gamification research concerns evaluating the impact of diverse game elements in varying contexts and platforms (Jia et al. 2017; Kumar 2013; Mekler et al. 2013b).

A diverse set of studies have considered the role that gamification and games play in motivating participation in a number of human computation games (Bowser et al. 2013; Eveleigh et al. 2013; Iacovides et al. 2013). Results from such findings have been mixed, with players identifying both positive and negative aspects of gamification in VCS. However, these studies draw on survey-based methodologies and predominantly focus on qualitative measures of volunteer participation, while quantitative analyses have suggested that self-reported motivations do not align with volunteer contributions (Mekler et al. 2013a). Our paper aims to contribute to this area of research by providing quantiative analysis of the impact of competitions on particpation.

We further note a number of studies and findings unrelated to VCS within the wider literature of relevance to our analysis, which we seek to build on to further the understanding of gamification on engagement and in HCI more broadly. Jia et al. explored the impact of individual gamification elements on participants according to personality characteristics (Jia et al. 2016). In team-based task completion, partnering with highly effective teammates has been found to reduce players own intrinsic motivations and perceptions of tasks (Luu and Narayan 2017). Brouwer conducted a survey of professional work teams in the Netherlands around the subject of intra-team competition and found that such competition has both negative and positive effects on team performance, through increased task complexity but reduced psychological safety (Brouwer 2016).

More specific to Games With A Purpose, but in a non-citizen science context, Siu et al. explored the impact of

[1] https://EyeWire.org/explore

competition-based scoring and reward mechanisms compared with collaboration-based reward mechanisms (Siu, Zook, and Riedl 2014). In particular, the authors found that intrinsic motivation and associated contest-design decisions were key to driving participation in competitions, but nevertheless must be balanced with extrinsic factors. In a similar domain to our research, Zheng et al. analysed motivations for participating in crowdsourcing competitions in paid microtask work, showing that engagement is linked to the nature of contest demands (Zheng, Li, and Hou 2011). Similarly, introducing cognitively demanding challenges to games has been found to increase player activity, while physically challenging tasks do not (Cox et al. 2012). Our work seeks to build on these studies by exploring the impact of adding temporary competitions to a specifically VCS context, where both altruistic and intrinsic motivations drive participation.

EyeWire

EyeWire is a crowdsourced citizen science Game With A Purpose, where participants contribute to identifying the presence of neuron cells in Magnetic Resonance images of mouse retina, mapping each image for further research (Kim et al. 2014; Tinati et al. 2017). Players do so by ‘tracing’ a three-dimensional section of an imaged neuron cell by manipulating two-dimensional cross-sectional images (see Figure 1). These 3D segments are known as cubes and it is these cubes that players submit to the project. Each cube starts with a ‘seed’ generated algorithmically and players must trace the path of the neuron from this seed through the cube. Multiple players trace each cube to generate an overall consensus as a means to improve the accuracy of submissions (Kim et al. 2014).

In order to convert this relatively repetitive task into an engaging game, EyeWire draws on a number of features. Players earn points for each cube traced, changing position within an integrated leaderboard, displaying daily, weekly and monthly points earned by each contributing user (Reeves et al. 2017; Tinati et al. 2017). The EyeWire science team regularly organises competitions, where players contribute through the main EyeWire interface as a team or as individuals (Tinati et al. 2017). Less frequently, the team organises large scale events, where a number of competition categories occur within a week-long period, united by a common storyline[2].

As well as the scientific task, EyeWire features a variety of discussion platforms including an integrated instant messenger chat service, forum and Wiki. Studies of chat participation have revealed that the instant messenger chat is vital for knowledge sharing, allowing new participants to learn from experienced players, promoting collaboration on tasks and encouraging players to continue to participate in EyeWire (Tinati, Simperl, and Luczak-Roesch 2017) and the chat interface is heavily used by a small but very active proportion of players (Tinati et al. 2015a).

The EyeWire task is cognitively complex and demands a significant time investment from participants to master the game. Rather than simply gathering or analysing data, volunteers in EyeWire must solve problems: identifying unpredictable branching neuron pathways, which span thousands of cubes (Lieberoth et al. 2014; Kim et al. 2014). EyeWire does not require participants to possess domain expertise or previous training (Borrett and Hughes 2016; Tinati et al. 2015a) but barriers for successful participation in the project are relatively high as players must first develop relevant skills (Lieberoth et al. 2014). Although the game provides a short compulsory tutorial, the most successful and accurate contributors have invested “tens of hours” over the course of “thousands of cubes” (Kim et al. 2014).

[2]see: http://blog.eyewire.org/competitions-the-master-guide/

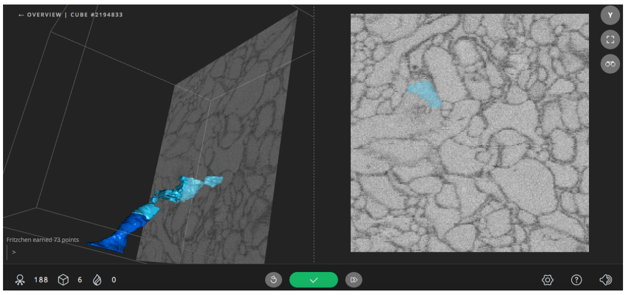


Figure 1: EyeWire tracing interface, showing left hand 3D cube view with seed and tracing, right hand 2D cross-sectional view and bottom left corner live chat overlay.

Data and Methods

To begin our analysis, we first sought to understand the role of competitions in the EyeWire project, as well as the needs of the project scientists and conditions for project success. To this end, we conducted a series of interviews with six members of the EyeWire team, over the course of two days. These participants included the project director, three ‘game masters’ (responsible for the day to day management of the community and monitoring manual and automated game processes), one designer and one developer. The six individuals represent the whole of the full-time EyeWire team and offered a diverse range of specialist knowledge about the history and outcomes of the project.

Following an approach used in a similar study (Tinati et al. 2015b), interviews used a structured reflection approach. During each session, participants were encouraged to share views, observations and experiences from the lifetime of the project. To guide participants and to provide a starting point for deeper discussions, we provided a series of questions framed around three key themes, including design decisions resulting in the use of competitions, expected and desired behaviours and measures and metrics of project success. After transcribing each interview, two researchers individually coded each of the transcripts, using our initial framework of three themes. After this process was complete, we compared the coded transcripts and identified and corrected areas of disagreement. From these coded transcripts, we derived a set of hypothesised participation patterns based on the experiences of the project team, which we identify as H.1 - H.6 (see *Interview Results - Desired Outcomes*).

Using the hypotheses outlined through the structured reflection process, we collected datasets from the EyeWire platform for quantitative analyses aimed at confirming the validity of each hypothesis. An initial dataset identified each player who contributed to Eyewire, as well as their number of contributions (in terms of points and cubes) for each 24 hour period from August 2016 - August 2017, including a separate log of final scores for each competition during this period. In response to the chat-related hypotheses put forward by the team, we also gathered a second dataset over a (91 day) period from June - August 2017. This dataset consisted of a time-stamped datastream containing discussion messages, achievements and system messages for all users active during these 91 days. Since players can contribute to the *Trivia* class through chat commands and the use of a bot set up for this competition class, we were unable to completely access data for this competition class. As a result and given the lack of hypotheses regarding this competition class, we removed all visible trivia entries from the chat data before further analysis.

For each dataset, we calculated relevant descriptive statistics, including mean total player numbers, cube contribution numbers, chat participant and chat message counts. Since competition participation in EyeWire is entirely voluntary, we distinguish between *competitors* who choose to participate in a given competition and *non-competitors* who do not. For each hypothesis (H1-H6) we defined suitable comparative test conditions. Using these conditions, we selected all available data from the relevant dataset - i.e, for *H1*, using dataset one we selected player numbers for all days on which competitions were held and compared these with all days on which they were not. Given the non-normal distributions and unpaired nature of the datasets involved, we used the non-parametric Mann-Whitney U test to evaluate each hypothesis. We report these test statistics using the Z-test, for ease of interpreation. A further set of additional tests (A1 - A4) were carried out to explore the impact of competitions on non-competing players. In addition to test statistics, we calculated a measure of effect size, Cohen's *d*, for each hypothesis, which we interpret using Sawilowsky's revisions of Cohen's original recommendations (Sawilowsky 2009). A summary of tests carried out, statistics and findings can be seen in table 3.

Results - Interviews

Design Requirements

The most significant factor influencing the implementation of regular competitions in the EyeWire project was to overcome difficulties arising from the nature of VCS games. During the interviews, the team described a range of characteristics they associated with successful and engaging games, which they had struggled to introduce to EyeWire. As a VCS game, they noted that there was no end goal at which point players could feel they had beaten the game - no final boss to defeat or challenge to overcome.

> "The ultimate reward systems, in some games, is to give the feeling of having beaten something... Almost no citizen science game is going to give that feeling."

While the scientific research aims associated with VCS serve as end goals, the team noted that research moves far slower than many players had first assumed. As a result, in early interactions between players and researchers, the players became disheartened at the slow rate with which research was progressing.

> "One of the players was like 'so what have you learned from our neurons since last time'? And I think one of the researchers was like... 'Oh, we haven't even looked at them'... It made the player feel like crap, because 'oh man, we've been working so hard!'... So in a way, to ease the burn of the slowness."

Furthermore the highly challenging and potentially abstract nature of some final research goals (for example "curing cancer") makes them unsuitable for introduction as final 'end point' challenges. Due to this, the team were unable to introduce more diegetic game elements such as narratives and characters - a factor they felt lead to further issues, such as a lack of progression.

> "One thing that's important for some games is to know you're getting better at it... [In Eyewire] there isn't a level system. You aren't playing level 20 because there is no level 20."

In essence, they were unsure how to make the game 'fun'. These difficulties were highlighted during the launch of the game, when despite initial interest in the project, contribution levels quickly fell.

> "Because we launched Eyewire, there was a huge bump in traffic and then after like, a month it just dwindled back down and we were like 'crap, what's going on?'"

The team concluded that this was due to the relatively basic nature of the game and a lack of gaming and social elements. At the time, the Eyewire interface was relatively basic. Players had access to the live chat function, cube tracing tool and a single, simplified leaderboard which reset at the end of the day, through which players could track their point scores.

Looking at the areas they felt to be strengths of citizen science games, the team made the decision to leverage the social and community aspects of the game. From this, they identified competition and associated events as a key design area that was lacking from Eyewire. Although the game had some competitive elements such as points and leaderboards, the team felt that these were insufficient for attracting participation without competition to contextualise and offer meaning to point scores.

> "If Mario were going along just collecting coins it would sort of be like 'woo, he's getting coins? Who cares! The coins don't... translate into anything!"

More specifically, given the lack of opportunities for overarching narratives and progression, these competitive elements were to form the basis of converting the Eyewire task into a game:

> "There were so many elements of Eyewire that were missing. There was no competition... You know, a game

Table 1: Competition Type.

Type	Description	Availability
Accuracy Happy Hour	Achieve accuracy above 80%	Major
Evil Cubes	Complete 12 very difficult cubes	Major
Happy Hour	Receive bonus for performance over two hours	Scheduled, Special
Marathon	Collaboratively complete cell within 24 hours	Minor, Major
Team Versus	Earn most points within time limit as team	Minor, Major
Trivia	Answer chat questions quickly and accurately	Major

[like Eyewire] without competition is arguably not a game. We had a leaderboard, but that leaderboard reset every day."

Leaderboards were expanded to allow players to track their performance over periods of days, weeks and months relative to other players. More significantly, however, players were invited to take part in ad-hoc competition challenges, where individual high-performing players would challenge the community to identify the best EyeWire tracer. Given the positive response to these contests, the team expanded the competitions, introducing regular team-based point generating challenges and regular collaborative timed cell completion challenges. Ultimately, so-called *major* competitions were introduced: week long events during which different competition challenges are held within an overarching diegetic narrative. Players must complete competition challenges to advance the narrative and overcome obstacles - for example, receiving clues to the identity of a killer in a whodunnit murder mystery themed event.

In contrast with minor competitions, major competitions are held less regularly (quarterly rather than bi-weekly/monthly) and offer specific rewards such as prizes and limited edition badges which players can display on their player profile. While minor competitions have a theme (for example, analogue vs digital) this theme expands only to player team names and a blog posting accompanied by artwork. Major competitions receive fully fleshed out narratives, with blog artwork and characters and in-game thematic elements such as in character chat communication from game masters. There is no difference in the core gameplay of minor and major competition variants, or in how long the individual competition is held for. A summary of competition types can be seen in Table 1.

Desired outcomes

Unlike some other citizen science projects, where existing work is carried out by algorithms (e.g., (Khatib et al. 2011)) or small science teams (e.g., (Lintott et al. 2008)), the research team predominantly relied on paid professionals to trace neurons prior to Eyewire. The success of EyeWire is therefore predominantly judged in terms of cost-effectivenss and efficiency - that is, is it more efficient to gather data through the large-scale crowdsourcing process rather than paying individual professionals? Initially the team found it difficult to improve on the performance of the paid professionals, noting three key issues. The first is that the tracing task is complex and takes time for individuals to learn to perform traces at a high enough level of accuracy for use in research. Two further concerns relate to the number of contributors and the efficiency with which players contribute to EyeWire. As with other citizen science platforms, a small minority of players do the majority of the work (Ponciano et al. 2014; Tinati et al. 2015a).

In the case of EyeWire, however, a minority of specially selected top-performing players must not only complete tracing tasks, but also administrative tasks to maintain site functionality platform. Dubbed 'scythes', these players are tasked with correcting inaccuracies in player traces and marking cubes as either completed or in need of further work. Without this work, the underlying algorithms may generate seeds for players, known as 'mergers', leading to wasted trace effort. Yet despite dedicating large volumes of time to the platform, even these highly active users are also fairly inefficient, dedicating a majority of their time to the IM chat interface rather than tracing (Tinati et al. 2015a).

Given these challenges, the first outcome of introducing competitions anticipated by the team was an increase in the number of EyeWire players contributing cubes to the project. In particular, the team have focussed their efforts on major competitions which feature limited edition 'badge' rewards which players can display in their in-game profile, occasional EyeWire branded merchandise prizes and the use of e-mail to notify players, particularly lapsed and inactive players - those who had been inactive for months or longer - of upcoming major competitions. From these we derive two hypotheses:

H.1 - Competitions lead to an increased number of participants compared with non-competition periods.

H.2 - Major competitions attract more players than their minor counterparts.

A second and perhaps unsurprising desired outcome is to increase overall effiency and the number of cubes submitted by players. While the team were generally optimistic about the effect of competition activities on cube contribution rates, they nevertheless noted the existence of a group of 'lurkers' - a proportion of the EyeWire player base who contribute in a steady manner and are unaffected by competitions or other efforts to increase their participation in the game. For this reason, the team stated their opinions that competition effectiveness would be largely dependent on whether players chose to participate or not.

H.3 - Competitions lead to increased cube contributions over non-competition periods.

H.4 - Levels of cube contribution from non-participating players are unaffected by competitions.

The third outcome described by the participants was to increase participation within the IM chat interface. When introducing competitive elements, the team also strengthened social aspects of the project to create an engaging game atmosphere. The most important of these elements was use of the IM chat interface. The team described two key aims of chat participation: increasing participation in EyeWire by providing an interruption from task completion which could otherwise prove monotonous and improving player performance by answering queries and through specially selected

Table 2: Daily mean statistics for Each Competition class (Dataset 1) * - Chat statistics unavailable for happy hour extra class due to a lack of instances of this class.

Type	Players	Percent Competing	Cubes	Chat Participants	Chat Messages	Percent of messages from competitors
Acc HH	238.67	25.24%	5395.83	78.50	1606.00	70.21%
Happy (Sched)	153.61	34.94%	4179.94	41.57	746.35	78.29%
Happy (Extra)	133.15	38.44%	3971.08	N/A*	N/A*	N/A*
Marathon Min	146.75	35.05%	6300.75	37	766	87.86%
Marathon Maj	164.6	40.74%	5996.63	31	506	75.60%
Team Vs Min	152.59	36.58%	4429.09	45.36	955.20	77.75%
Team Vs Maj	191.20	29.05%	5887.38	43.80	911.40	81.37%
None	139.18	N/A	3503.75	27.07	360.61	N/A

player 'mentors' with access to tools to help new and inexperienced players. Major competitions make use of use the chat interface to provide thematic hints and to hold the *trivia* competition type, which is unique to major competition periods. However, since these aspects are absent from minor competitions, the team were sceptical that chat participation would increase substantially outside of major periods.

H.5 - Only major competition periods generate increased numbers of chat messages.

H.6 - Only major competition periods lead to increased numbers of players participating in chat.

Statistical Analysis

Player Numbers In contrast to the expectations of the Eyewire team, only a small minority of EyeWire players directly participate in competition events. Within a one year period, a total of 10,296 players contributed at least 1 cube to the EyeWire project. However, only 494 players earned at least one point, or contributed at least one cube to one of the 143 competitions held during this time, a figure which represents just 4.80% of the active players during this year. A summary of player statistics for the different competition types can be seen in Table 2. With the exception of the "extra" happy hour class, the mean number of players is higher for all competition days than for days on which competitions are not held. It should be noted that although the Accuracy Happy Hour appears to be the most popular competition, we believe that this is largely due to the day on which it is held, which is always the first day of the major competitive period. This may explain the large number of active players, as far from being a highly popular competition, the accuracy happy hour attracts on average 13 players.

Competitions nevertheless have a significant effect on the number of active players in Eyewire. On days when competitions are held, player numbers are moderately higher than non-competition days (d=0.55, $p<0.01$). Major competitions are also seen to be much more effective in attracting players than minor competitions (d=0.88, $p<0.01$), in line with H.2 as suggested by the Eyewire team. In fact, our analysis shows that this is predominantly due to a proportion of players who do not participate in competitions yet choose to sign in and contribute on days when competitions are held. Not only is there a statistically significant increase in active players from the non-competing group on competition days (A1, d=0.55, $p<0.01$), but this effect is greater during major competition periods over minor periods (A2, d=0.92, $p<0.01$). On the other hand, although these increases are statistically significant, these represent a small minority of the 9,802 players who chose not to join any competitions.

Task Completion Players who sign up for competitions complete more work than players who are otherwise active on the platform. Competitions have a medium to large effect on the number of cubes contributed by competition participants as a group (d=0.79, $p<0.01$), as well as a small to medium effect on the number of cubes per competing player (d=0.49, $p<0.01$) as can be seen in figure 2. Nevertheless, further analysis reveals that this increase in cube contributions extends to non-participating players, who as a group display a small increase in the number of cubes submitted on days where competitions are held (A3, d=0.35, $p<0.01$), particularly on major competition days (A4, d=0.83, $p<0.01$). In terms of competition types, cube contributions vary greatly. While all competition types encourage greater numbers of cube submissions, the Marathon class attracts higher numbers of cubes in comparison with other classes, particularly the minor marathon class.

Community Discussion An analysis of the 91 day sample of chat participation reveals a notable increase in chat participation during competitions. When comparing days on which competitions occur with those on which no competitions are held, test statistics identify a moderate increase in the total number of chat participants (d=0.60, $p<0.01$). Although the majority of the mechanisms introduced by the Eyewire team to motivate chat participation were focused on major competitive periods, both major and minor competitions demonstrate significant increases in the numbers of chat messages and numbers of messages sent per player ($p<0.01$). Nevertheless, this effect is larger for major competitions (d=0.85) than for minor competitions (d=0.62). However, when comparing the number of chat participants on minor and major competition days with non-competition days, the analysis was inconclusive.

Discussion

Our results show that while competitions appeal to a minority of human computation contributors, they nevertheless have a significant positive impact on the number of contributions generated, as well as the level of community discussion activity. This effect is in part due to the effect that competitions have on those participants who choose not to actively compete, yet who nonetheless choose to sign-in and contribute more during competition periods. Our findings suggest that competitions have an overall positive effect on participation in both chat and task activities. We note, however, that EyeWire is a highly gamified context and suggest that further analysis is required to understand the use of competitions in non-gamified human computation initiatives - particularly those which rely solely on altruistic and intrinsic motivations.

Table 3: Summary of Hypothesis, *H*, and additional Mann-Whitney U test, *A*, outcomes. Z-crit for p<0.01 = -2.33

H	Condition	Z-Stat	*d*	Conclusion
H1	Number of players (competition vs non-competition)	-5.1	0.55	Hypothesis confirmed (p<0.01, medium effect size)
H2	Number of players (minor vs major)	-3.5	0.88	Hypothesis confirmed (p<0.01, large effect size)
H3	Total cube contributions (competition vs non-competition)	-6.99	0.79	Hypothesis confirmed (p<0.01, medium/large effect size)
	Cube contributions per player (competitions vs non-competition)	-4.58	0.49	Hypothesis confirmed (p<0.01, small/medium effect size)
H4	Total cube contributions for non-competing players (competitions vs non-competition)	-3.33	0.35	Hypothesis incorrect. Cube levels from non-competing players are higher during competitions (p<0.01, small effect size)
H5	Chat messages per player (major vs non-competition)	-3.08	0.85	Hypothesis incorrect. Both minor and major competitions generate increased chat messages compared with non-competition days (p<0.01, large, medium effect size)
	Chat messages per player (minor vs non-competition)	-2.37	0.62	
H6	Total chat participants (competitions vs non-competition)	-2.64	0.60	Uncertain. Results show an increase in chat participants over all competition types (p¡0.01, medium effect size)
A1	Numbers of non-competing players (competition vs non-competition)	-5.1	0.55	Competitions result in a significant increase in non-competing player numbers (p<0.01, medium effect size)
A2	Numbers of non-participating players (minor competitions vs major competitions)	-3.65	0.92	Major competitions attract significantly more non-competing players than minor competitions (p<0.01, large effect size)
A3	Total cube contributions from non-competing players (competitions vs non-competitions)	-3.33	0.35	During competitions, cube contributions from non-competing players increase (p<0.01, small effect size)
A4	Total cube contributions from non-competing players (minor competitions vs major competitions)	-3.36	0.83	During major contributions, more cubes are contributed by non-competing players than during minor competitions (p<0.01, large effect size)

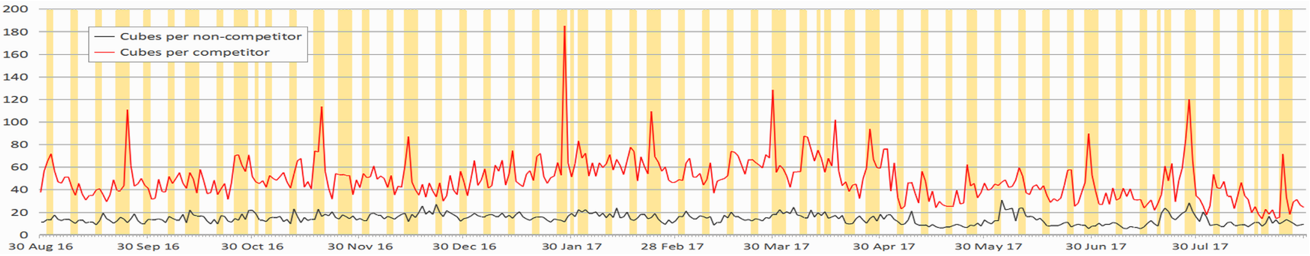


Figure 2: Cubes per non-competing and competing players on competition (highlighted) and non-competition days.

Major competitions attract significantly more competitors, non-competing players and cube contributions than minor competition events. Although the team predicted an increase in competitor and cube numbers, it is nevertheless surprising that major competitions should be significantly more effective than minor competitions at encouraging cube contributions. While major events feature limited edition badge rewards for participation, these are given to all players who reach the (relatively low) threshold of points or cubes - dependent on competition type - and there are no other mechanisms to encourage or reward contributions that cannot also be earned in minor competitions. Previous studies have demonstrated that participants in human computation initiatives are more motivated and engaged by diegetic, narrative and progression-driven gamified experiences, even when the underlying task itself is identical (Prestopnik and Tang 2015). Our findings suggest that this may be true even in circumstances where the elements are more implicit and community-driven - that is, in the case of EyeWire, where the only visible elements of the progression and narrative are visible through the chat or blog features, such as ’whodunnit’ clues.

Similarly, although our findings demonstrate an increase in non-competitor numbers and participation levels during competitions, the reason for this is unclear. EyeWire features few opportunities for spectation or collaboration with other players and competition rewards are provided only to those who participate. One possibility, as raised by the team, is the potential impact of a positive chat atmosphere - as the team stated: *“chat is blowing up. Everyone’s excited.”* Moreover, it should be noted that the chat and community-driven narrative features are still visible to these non-competing players, but only during major competitions - a factor that would explain the significant increase in non-competitor numbers and contributions during the major competition class. We believe this is an interesting area for further study and that the inclusion of more visible competition elements that include the whole community may be important for maximising the effectiveness of competitions in human computation initiatives. A further possibility is that the increase in non-competitor numbers is a result of notification campaigns (particularly e-mails) attached to competitions - particularly since this effect is greater during major competitions, where such campaigns are held. Our data was insufficient to test such a hypothesis, as we lack confirmation that players had received or seen campaign emails (typically sent shortly before competitions occur) and whether notifications or competitions - or both - drive any return to EyeWire. Nevertheless, our findings show that any impact on participation is extremely brief, coinciding only with competitions (see figure 2). Moreover, this does not explain the increase in participation levels from within this group during competitions.

Within the literature, autonomy has been highlighted as a key motivational affordance across a variety of human computation initiatives (Lukyanenko, Parsons, and Wiersma 2016; Prestopnik and Tang 2015). However, our findings note a potential trade-off with regard to autonomy in competitions. The team versus competition class - during which participants can earn points from any cube and, where applicable, bonuses, tutorials and promotion-dependent activities - attracted a greater number of competitors, but on average resulted in reduced numbers of cube contributions when compared with the marathon class. While the marathon class is significantly less autonomous, accepting only cube contributions and only for one cell, this class has a collaborative, goal-oriented aspect which is not present in the team versus competition. Rather than earning points relative to the performance of other players, players are instead scored based on their own number of cubes contributed and must as a community reach the required number of cells within the time-limit. We note that the effectiveness of competitive and collaborative elements remains a controversial topic within the wider literature and we suggest that further exploration is required to understand how autonomy fits within this space.

Further related to autonomy, a number of VCS studies have made explicit or implicit connections between autonomy and the chat platforms, suggesting that chat participation can replace or overcome a lack of autonomy in tasks (Jackson et al. 2014; Reeves et al. 2017). This view was even put forward by the EyeWire team: *"chat makes up for the dryness of the task."* On the other hand, our results suggest that this may not be the case, at least in EyeWire. During the less autonomous marathon class, fewer chat messages were sent by the community and the number of chat participants was lower than during the much more autonomous team versus category. This suggests that while more intensive, time-limited challenge activities may be effective for motivating task completion, these restrictions discourage autonomous activity and engagement with other tasks such as chat.

Finally, a key area for research, which we have not addressed in our analysis, is consideration of the impact of competitions on the quality and accuracy of responses. The nature of EyeWire makes it unsuitable for such analysis, as the platform largely lacks ground truth cubes, assesses responses through majority voting and includes an iterative process through which cubes can be corrected and marked for further analysis by the promoted 'scythe' player class (Reeves et al. 2017). In contrast with the view held by the team that competitions can effectively train players to contribute to human computation tasks, our findings demonstrate that the accuracy happy hour class attracts only a small number of participants. However, this does not discount the possibility of players learning through competitions, as it takes many hours and thousands of cubes to 'master' EyeWire and as such, the increase in activity that competitions encourage is crucial for newer and less active players.

In addition, further consideration must be given to additional methods for overcoming the weaknesses and challenges of VCS and associated game genres. This may include testing other progression formats - for example, the EyeWire team suggested opportunities posed by badges or unlockables, offering individual-centred rather than communal progression. Opportunities for team or group play are also of interest, given the apparent importance of sociality in VCS, but we caution that such opportunities also pose issues for common quality assurance processes such as redundancy (Wiggins et al. 2011). Given the uncommon and temporary nature of competitions across VCS (Reeves et al. 2017), it is unclear to what extent our findings generalise, particularly to less socially-driven or gamified projects where participant motivations differ (Mekler et al. 2013b).

Limitations

Due to the nature of EyeWire's chat system and volume of messages produced, these messages are not readily archived and so we were only able to collect chat messages live from the system. For this reason, our second dataset covered only 1/4 of the larger first dataset. We noted nonetheless that this period covered 48 competition days, including a range of minor and major events and therefore felt it was suitable for analysis. To further overcome this issue, we analysed hypotheses covering minor and major competitions, rather than specific competition types, to maximise the volume of available data. We do, however, make the assumption that non-, minor and major competition activity during this 91 day period is indicative of activity in other periods.

Conclusion

Overall, this study contributes to the understanding of how competitions influence player participation in human computation activities. We have identified the underlying rationale for introducing competitions to Eyewire. Our findings demonstrate that competitions effectively increase levels of contribution for task and chat components of the project, while increasing the number of active players. This effect persisted even with those players who choose not to contribute to competitions, who exhibit greater numbers of active players and cube contributions on competition days. Based on these findings, we identified points of contention with previous human computation literature, as well as a number of areas for further research and analysis.

Acknowledgements

This research was funded by the Digital Economy programme, as part of UK Research and Innovation under grant agreement EP/G036926/1 and the Horizon2020 programme, through the Stars4All project under grant agreement 688135. We would like to thank the EyeWire team for their time and support and the anonymous reviewers for their comments and recommendations.

References

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

Borrett, S., and Hughes, L. 2016. Reporting methods for processing and analysis of data from serial block face scanning electron microscopy. *Journal of microscopy* 263(1):3–9.

Bowser, A.; Hansen, D.; He, Y.; Boston, C.; Reid, M.; Gunnell, L.; and Preece, J. 2013. Using gamification to inspire new citizen science volunteers. In *Proceedings of the first international conference on gameful design, research, and applications*, 18–25. ACM.

Brouwer, R. 2016. When competition is the loser: the indirect effect of intra-team competition on team performance through task complexity, team conflict and psychological safety. In *System Sciences (HICSS), 2016 49th Hawaii International Conference on*, 1348–1357. IEEE.

Cheng, J.; Teevan, J.; Iqbal, S. T.; and Bernstein, M. S. 2015. Break it down: A comparison of macro-and microtasks. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, 4061–4064. ACM.

Cox, A.; Cairns, P.; Shah, P.; and Carroll, M. 2012. Not doing but thinking: the role of challenge in the gaming experience. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 79–88. ACM.

Cox, J.; Oh, E. Y.; Simmons, B.; Graham, G.; Greenhill, A.; Lintott, C.; Masters, K.; et al. 2015. Doing good online: An investigation into the characteristics and motivations of digital volunteers.

Curtis, V. 2014. Online citizen science games: opportunities for the biological sciences. *Applied & translational genomics* 3(4):90–94.

Deterding, S.; Dixon, D.; Khaled, R.; and Nacke, L. 2011. From game design elements to gamefulness: defining gamification. In *Proceedings of the 15th international academic MindTrek conference: Envisioning future media environments*, 9–15. ACM.

Dickinson, J. L.; Zuckerberg, B.; and Bonter, D. N. 2010. Citizen science as an ecological research tool: challenges and benefits. *Annual review of ecology, evolution, and systematics* 41:149–172.

Eveleigh, A.; Jennett, C.; Lynn, S.; and Cox, A. L. 2013. "i want to be a captain! i want to be a captain!": Gamification in the old weather citizen science project. In *Proceedings of the first international conference on gameful design, research, and applications*, 79–82. ACM.

Greenhill, A.; Holmes, K.; Woodcock, J.; Lintott, C.; Simmons, B. D.; Graham, G.; Cox, J.; Oh, E. Y.; Masters, K.; and Lewandowski, D. 2016. Playing with science: exploring how game activity motivates users participation on an online citizen science platform. *Aslib Journal of Information Management* 68(3).

Haklay, M. 2013. Citizen science and volunteered geographic information: Overview and typology of participation. In *Crowdsourcing geographic knowledge*. Springer. 105–122.

Iacovides, I.; Jennett, C.; Cornish-Trestrail, C.; and Cox, A. L. 2013. Do games attract or sustain engagement in citizen science?: a study of volunteer motivations. In *CHI'13 Extended Abstracts on Human Factors in Computing Systems*, 1101–1106. ACM.

Jackson, C.; Østerlund, C.; Crowston, K.; Mugar, G.; and Hassman, K. 2014. Motivations for sustained participation in citizen science: Case studies on the role of talk. In *17th ACM Conference on Computer Supported Cooperative Work & Social Computing*. Citeseer.

Jennings, N. R.; Moreau, L.; Nicholson, D.; Ramchurn, S.; Roberts, S.; Rodden, T.; and Rogers, A. 2014. Human-agent collectives. *Communications of the ACM* 57(12):80–88.

Jia, Y.; Xu, B.; Karanam, Y.; and Voida, S. 2016. Personality-targeted gamification: a survey study on personality traits and motivational affordances. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, 2001–2013. ACM.

Jia, Y.; Liu, Y.; Yu, X.; and Voida, S. 2017. Designing leaderboards for gamification: Perceived differences based on user ranking, application domain, and personality traits. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, 1949–1960. ACM.

Khatib, F.; Cooper, S.; Tyka, M. D.; Xu, K.; Makedon, I.; Popović, Z.; Baker, D.; and Players, F. 2011. Algorithm discovery by protein folding game players. *Proceedings of the National Academy of Sciences* 108(47):18949–18953.

Kim, J. S.; Greene, M. J.; Zlateski, A.; Lee, K.; Richardson, M.; Turaga, S. C.; Purcaro, M.; Balkam, M.; Robinson, A.; Behabadi, B. F.; et al. 2014. Space-time wiring specificity supports direction selectivity in the retina. *Nature* 509(7500):331.

Kumar, J. 2013. Gamification at work: Designing engaging business software. In *International Conference of Design, User Experience, and Usability*, 528–537. Springer.

Lieberoth, A.; Pedersen, M. K.; Marin, A. C.; Planke, T.; and Sherson, J. F. 2014. Getting humans to do quantum optimization-user acquisition, engagement and early results from the citizen cyberscience game quantum moves. *Human Computation* 1(2).

Lieberoth, A. 2015. Shallow gamification: Testing psychological effects of framing an activity as a game. *Games and Culture* 10(3):229–248.

Lintott, C. J.; Schawinski, K.; Slosar, A.; Land, K.; Bamford, S.; Thomas, D.; Raddick, M. J.; Nichol, R. C.; Szalay, A.; Andreescu, D.; et al. 2008. Galaxy zoo: morphologies derived from visual inspection of galaxies from the sloan digital sky survey. *Monthly Notices of the Royal Astronomical Society* 389(3):1179–1189.

Luczak-Roesch, M.; Tinati, R.; Simperl, E.; Van Kleek, M.; Shadbolt, N.; and Simpson, R. 2014. Why won't aliens talk to us? content and community dynamics in online citizen science.

Lukyanenko, R.; Parsons, J.; and Wiersma, Y. F. 2016. The impact of task design on accuracy, completeness and discovery in surveillance-based crowdsourcing. In *Proceedings of the 4th International Conference on Collective Intelligence*.

Luu, S., and Narayan, A. 2017. Games at work: Examining a model of team effectiveness in an interdependent gaming task. *Computers in Human Behavior*.

Mattos, A. B.; Herrmann, R.; Shigeno, K. K.; and Feris, R. 2014. A mission-oriented citizen science platform for efficient flower classification based on combination of feature descriptors. In *EMR@ ICMR*, 45–52.

Mekler, E. D.; Brühlmann, F.; Opwis, K.; and Tuch, A. N. 2013a. Disassembling gamification: the effects of points and meaning on user motivation and performance. In *CHI'13 extended abstracts on human factors in computing systems*, 1137–1142. ACM.

Mekler, E. D.; Brühlmann, F.; Opwis, K.; and Tuch, A. N. 2013b. Do points, levels and leaderboards harm intrinsic motivation?: an empirical analysis of common gamification elements. In *Proceedings of the First International Conference on gameful design, research, and applications*, 66–73. ACM.

Michelucci, P. 2013. *Handbook of Human Computation*. Springer Science & Business Media.

Oliveira, N.; Jun, E.; and Reinecke, K. 2017. Citizen science opportunities in volunteer-based online experiments. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, 6800–6812. ACM.

Østerlund, C.; Mugar, G.; Jackson, C. B.; Hassman, K.; and Crowston, K. 2014. Socializing the crowd: Learning to talk in citizen science. In *Academy of Management Annual Meeting, OCIS Division, Philadelphia, PA*.

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

Ponti, M.; Hillman, T.; and Stankovic, I. 2015. Science and gamification: The odd couple? In *Proceedings of the 2015 Annual Symposium on Computer-Human Interaction in Play*, 679–684. ACM.

Prestopnik, N. R., and Tang, J. 2015. Points, stories, worlds, and diegesis: Comparing player experiences in two citizen science games. *Computers in Human Behavior* 52:492–506.

Prestopnik, N.; Crowston, K.; and Wang, J. 2014. Exploring data quality in games with a purpose. *iConference 2014 Proceedings*.

Reeves, N.; Tinati, R.; Zerr, S.; Van Kleek, M. G.; and Simperl, E. 2017. 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*, 2137–2152. ACM.

Rotman, D.; Preece, J.; Hammock, J.; Procita, K.; Hansen, D.; Parr, C.; Lewis, D.; and Jacobs, D. 2012. Dynamic changes in motivation in collaborative citizen-science projects. In *Proceedings of the ACM 2012 conference on computer supported cooperative work*, 217–226. ACM.

Sandbrook, C.; Adams, W. M.; and Monteferri, B. 2015. Digital games and biodiversity conservation. *Conservation Letters* 8(2):118–124.

Sauermann, H., and Franzoni, C. 2015. Crowd science user contribution patterns and their implications. *Proceedings of the National Academy of Sciences* 112(3):679–684.

Sawilowsky, S. S. 2009. New effect size rules of thumb. *Journal of Modern Applied Statistical Methods* 8(2):26.

Seaborn, K., and Fels, D. I. 2015. Gamification in theory and action: A survey. *International Journal of Human-Computer Studies* 74:14–31.

Siu, K.; Zook, A.; and Riedl, M. O. 2014. Collaboration versus competition: Design and evaluation of mechanics for games with a purpose. In *FDG*.

Sprinks, J.; Houghton, R.; Bamford, S.; and Morley, J. 2015a. Citizen scientists: The importance of being needed and not wasted. In *Proceedings of the 2015 Annual Symposium on Computer-Human Interaction in Play. ACM*.

Sprinks, J.; Houghton, R.; Bamford, S.; and Morley, J. 2015b. The impact of task workflow design on citizen science users and results. *Contemporary Ergonomics and Human Factors* 371–378.

Tinati, R.; Luczak-Rösch, M.; Simperl, E.; Shadbolt, N.; and Hall, W. 2015a. '/command'and conquer: Analysing discussion in a citizen science game. In *Proceedings of the ACM Web Science Conference*, 26. ACM.

Tinati, R.; Van Kleek, M.; Simperl, E.; Luczak-Rösch, M.; Simpson, R.; and Shadbolt, N. 2015b. Designing for citizen data analysis: a cross-sectional case study of a multi-domain citizen science platform. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, 4069–4078. ACM.

Tinati, R.; Luczak-Roesch, M.; Simperl, E.; and Hall, W. 2017. An investigation of player motivations in eyewire, a gamified citizen science project. *Computers in Human Behavior* 73:527–540.

Tinati, R.; Simperl, E.; and Luczak-Roesch, M. 2017. To help or hinder: Real-time chat in citizen science. *Proceedings of ICWSM* 2017.

Wiggins, A., and Crowston, K. 2011. From conservation to crowdsourcing: A typology of citizen science. In *System Sciences (HICSS), 2011 44th Hawaii international conference on*, 1–10. IEEE.

Wiggins, A.; Newman, G.; Stevenson, R. D.; and Crowston, K. 2011. Mechanisms for data quality and validation in citizen science. In *e-Science Workshops (eScienceW), 2011 IEEE Seventh International Conference on*, 14–19. IEEE.

Zheng, H.; Li, D.; and Hou, W. 2011. Task design, motivation, and participation in crowdsourcing contests. *International Journal of Electronic Commerce* 15(4):57–88.

Is Virtual Citizen Science A Game?

ELENA SIMPERL, University of Southampton
NEAL REEVES, University of Southampton
CHRIS PHETHEAN, University of Southampton
TODD LYNES, University of Southampton
RAMINE TINATI, University of Southampton

The use of game elements within virtual citizen science is increasingly common, promising to bring increased user activity, motivation and engagement to large-scale scientific projects. However there is an ongoing debate about whether or not gamifying systems such as these is actually an effective means by which to increase motivation and engagement in the long term. While gamification itself is receiving a large amount of attention, there has been little beyond individual studies to assess its suitability or success for citizen science; similarly, while frameworks exist for assessing citizen science performance, they tend to lack any appreciation of the effects that game elements might have had. We therefore review the literature to determine what the trends are regarding the performance of particular game elements or characteristics in citizen science, and survey existing projects to assess how popular different game features are. Investigating this phenomenon further, we then present the results of a series of interviews carried out with the EyeWire citizen science project team to understand more about how gamification elements are introduced, monitored and assessed in a live project. Our findings suggest that projects use a range of game elements with points and leaderboards the most popular, particularly in projects that describe themselves as 'games'. Currently, gamification appears to be effective in citizen science for maintaining engagement with existing communities, but shows limited impact for attracting new players.

CCS Concepts: • **Human-centered computing** → *Computer supported cooperative work*;

Additional Key Words and Phrases: Citizen science, Gamification, Social Computing, Engagement, Motivation

ACM Reference format:
Elena Simperl, Neal Reeves, Chris Phethean, Todd Lynes, and Ramine Tinati. 2018. Is Virtual Citizen Science A Game?. *ACM Trans. Soc. Comput.* , , Article (2018), 37 pages.
https://doi.org/

1 INTRODUCTION

Online, or virtual, citizen science (VCS) uses Web-based techniques to allow members of the public to participate as volunteers in a scientific research project, without possessing expert knowledge about the subject at hand. This may fall into almost any domain, with notable examples in astronomy, zoology, and healthcare where using the Web to facilitate citizen science has proved to be a successful application of crowdsourcing for problems which are computationally

We thank the EyeWire project team for their participation in the interviews for this research. This research was carried out thanks to the Stars4All EU Horizon 2020 Framework Programme grant, agreement no. 688135.

Manuscript submitted to ACM

unfeasible or inaccurate, such as classifying images of galaxies or identifying animals. However, while there are solid examples of success, a continuing challenge in VCS centers on how to increase both the number of participants, and their ongoing engagement with the project. This may be in order to create greater contributions with regards to the amount of work completed, the length of time the player participates, or the quality of the data that they produce by doing so, but also simply just to allow a wide range of players to contribute effectively, and in turn making the output of the project a useful contribution to science.

Drawing on many methods and theories from Human-Computer Interaction (HCI) and Computer-Supported Co-operative Work (CSCW), a key area of citizen science research looks at how a system can be designed in order to influence user behaviour and encourage people to interact with, and contribute to, the project. One such approach is to include game elements within the project, in order to increase the attractiveness of the work, motivate continued performance, and therefore improve user retention.

The inclusion of elements from game design within non-game environments, commonly described with the term 'gamification', has increased rapidly over recent years. This has occurred across a wide-range of applications and is a valuable tool for researchers and scientists wishing to undertake studies using large audiences [Deterding et al. 2011a]. Gamification has been studied previously in the context of VCS to determine how it can be used to increase participation (see, for example Tinati et al. [2016]); however while it has received a lot of attention, little has yet been done to look specifically at how the inclusion of specific elements relates to the subsequent success of the project.

There is also an ongoing debate about whether the use of extrinsic rewards, such as those offered through completing a game, are a good motivational strategy. In a review of 128 experiments, Deci et al. [1999] concludes that 'tangible' rewards (such as money) tend to have a negative effect on people's intrinsic motivation to work or contribute. According to Self-Determination Theory (SDT) humans seek out activities that satisfy motivational needs, one of which is autonomy, and the attachment of extrinsic motivators such as rewards can reduce this experience [Deci et al. 1999; Deterding 2011]. However as games are played voluntarily, there is an argument that this in itself provides the experience of autonomy, and thus the accumulation of points or badges for example through choosing to play the game itself is intrinsically motivating [Deterding 2011]. While there are certain conditions that can ensure that the effect of extrinsic rewards is negated, Deci et al. [1999] summarise their review by stating that focusing only on extrinsic rewards can risk a reduction in intrinsic motivation. In crowdsourcing, studies have shown that there a wide range of motivators, both intrinsic and extrinsic, behind why people choose to participate [Morschheuser et al. 2017]. In terms of citizen science games, McGonigal argues that the addition of 'meaning', or the sense that by playing the volunteer is belonging to and contributing something significant, is an intrinsic motivator that can make them feel a part of something much bigger than themselves [McGonigal 2011, p. 49-51]. This desire to participate in scientific research has also been shown to be one of the biggest draws to contributing to gamified projects, while the intellectual challenge as presented by the game then helped to maintain involvement [Curtis 2015], suggesting that this combination does not necessarily show the negative effects forecast by Deci, Koestner and Ryan in [Deci et al. 1999]. Previous VCS research, examining games such as FoldIt [Curtis 2015], EyeWire [Iacovides et al. 2013] and EteRNA [Bohannon 2014] tends to consider gamification to have a positive effect, but lacks in-depth analysis behind what the project teams mean by 'success' or how the use of game elements help to produce this. Further insights about the positive—or negative—impacts of gamification are required if there is to be an end to the gap in knowledge about how best to use these techniques.

The motivations of this paper are as follows. Firstly, there is a need to improve understanding around the success of game elements in citizen science projects and greater insights into how gamification can be and is used [Deterding et al. 2011a]. Previous research has been carried out regarding the level of success in citizen science in general (see, for example, the framework of success for assessing projects in Graham et al. [2015]). However, existing studies have–to the best of our knowledge–not looked broadly at how gamification might contribute to this success, such as which elements are effective in increasing user engagement, whether or not a gamified approach is typically successful, or how designers decide which game features to include. Further to this, VCS is a diverse field and there are currently few overviews that detail how game elements and characteristics are currently being used in this space — and equally crucially, why these design decisions are being made, particularly in cases where they diverge from literature recommendations.

To this end, we review findings from the VCS literature to assesse empirical findings on the reported performance of gamification across a range of VCS domains. We discuss the ways in which elements are implemented and review whether they have a positive, negative or mixed effect on volunteer engagement, performance and motivation. Based on these findings, we then carry out a survey of 31 VCS projects including games and non-games, to understand the extent to which findings and recommendations from the literature are put into practise and also to explore the distinction, if any, between what it means for a project to call itself a 'game'. To understand the reasoning behind these design decisions, we then supplement this literature and project survey with findings from interviews with all members of the design team behind the gamified VCS project Eyewire [1]. We find that certain game elements (points and leaderboards) are commonly used, with others such as badges, rewards and statuses building on these, and these elements are used more on projects that describe themselves as 'games'. Projects that do not describe themselves in this way appear to focus more on learning and education, suggesting a different approach to engaging volunteers. Nonetheless, these elements often diverge from those recommended in the literature, with extrinsic and competitive elements more commonly used than the cooperative and intrinsic, progression-based game play supported by the literature. Our interviews with the EyeWire team give insights into the reasoning behind such design decisions, showing that the project team are predominantly influenced by player feedback and the need to ensure long-term player engagement. Development restrictions play a crucial role in preventing the team from expanding the project and changes are typically made without much influence from what other citizen science games are reporting to have been successful. Throughout our findings we observe tensions between what it means to be a 'game' and the needs of a VCS project — particularly in terms of the expectations of players.

The remainder of this paper is structured as follows. First we present a background introduction to VCS and gamification, and how they have been used together. We then review the existing literature on the effects of gamification in VCS, examining studies to see how effective the use of game elements has been in real-world projects. The first contribution of this paper summarises findings from the literature identifying the impact of different game features on player engagement and output quality. Secondly, our survey of VCS projects explores the extent to which these recommendations are employed in projects and the popularity of different features both in contexts which use the 'game' descriptor and those which do not. Finally, our third contribution draws on the interview-based study with the EyeWire team to understand the decision making processes required for the successful design and implementation of VCS games. We envisage that these insights can be used by citizen science project designers to consider the potential impact of different game elements, along with how to track these effects so that meaningful insights can be gained from them, and by VCS scholars interested in tracking the overall success of gamification in this domain.

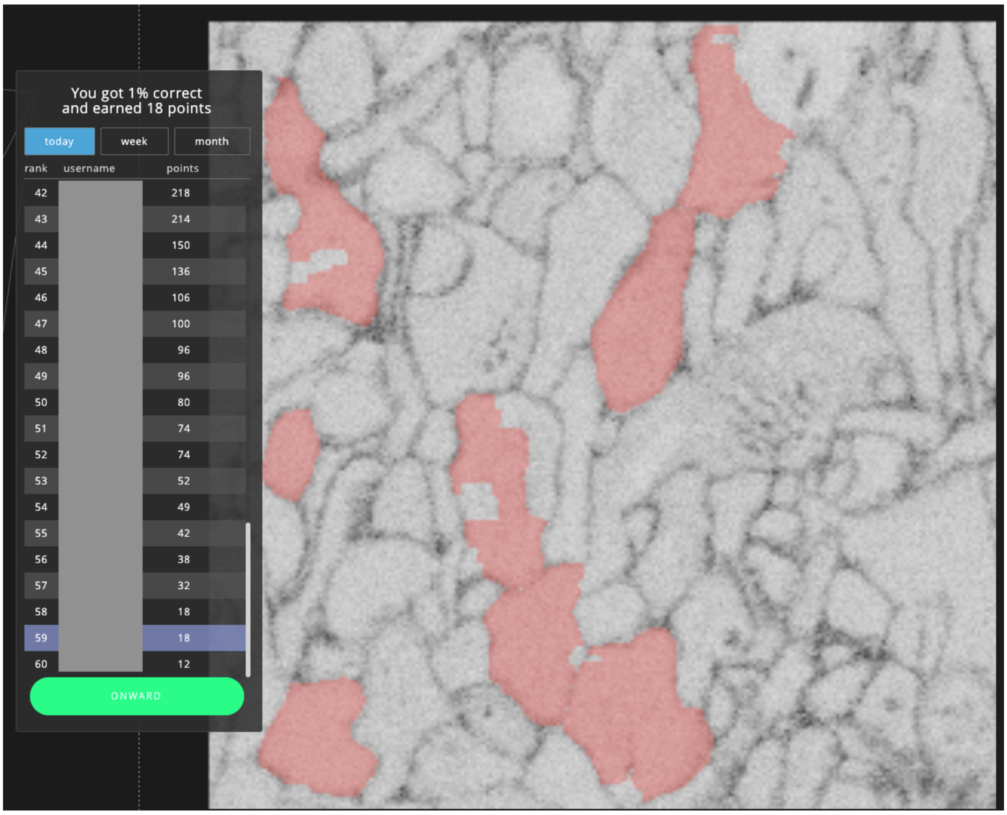


Fig. 1. The EyeWire game interface, showing accuracy, points acquired and the leaderboard. Usernames have been obscured for privacy. http://eyewire.org

2 BACKGROUND AND RELATED WORK

2.1 Virtual Citizen Science

VCS allows scientists to process large amounts of data using human computation, to complete tasks that would otherwise take unfeasible lengths of time or with potential sacrifices to accuracy to be carried out using machines. It draws upon theories from HCI and CSCW, and typically makes use of a digital platform to allow professional scientists to crowdsource contributions to their research from large numbers of people [Howe 2006]. The use of the public for collecting and analysing data is becoming an important part of science [Tinati et al. 2015] with projects showing notable levels of success—including for example the galaxy-classifying web application Galaxy Zoo [Lintott et al. 2008].

A VCS project can be based in almost any domain; current projects can be found in a hugely diverse range of topics. While citizen science projects can include manual, offline work such as the recording of wildlife in participants' gardens [2], VCS projects are typically conducted online (such as in EyeWire, Figure 1), making use of both the prevalence of the Web to reach huge numbers of people and of large computer clusters that can subsequently store, manage and

[1] http://eyewire.org/explore

[2] see, for example, the iNaturalist project, where participants track and record details of their wildlife observations https://www.inaturalist.org/

process the resulting deluge of data. The results are aggregated to provide a level of assurance that the data is accurate, with the aim being to allow the scientists to access a result at a much faster speed than if they had to process all the data themselves. For example, the crowd approach was shown to be more scalable and economical than an automated solution for analysing images related to particle-tracking problems [Chen et al. 2016].

Using crowdsourcing for science is not anything inherently new; services such as SETI@home [Anderson et al. 2002] have long offered participants the chances to allow processing time on their own computers to assist with scientific research. In recent years, however, increased focus has been placed on using the volunteers themselves as the means of computation—either in terms of data analysis or data collection and curation. The smartphone age has given rise to the opportunity for "participatory sensing", allowing volunteers to record all types of phenomena using their mobile device [Restuccia et al. 2016]. Relying on the ability of the human brain to complete tasks such as the analysis of images is quicker than even the most powerful of supercomputers [Kawrykow et al. 2012]. Similar to SETI@home, VCS projects allow people to contribute without being a professional scientist. Comparable to the process followed in an assembly line, a complex scientific task is split into a number of simpler 'microtasks', each of which is basic enough to be solved by a non-expert. The tasks may include cataloguing or classifying items, mapping structures within images, or transcribing text, following a set of guidelines. These guidelines mean that anyone can participate without the typical level of expertise required in the subject area [Tinati et al. 2015]. The EyeWire project, for example, breaks down a complex neuroscience problem of mapping the connectome in the brain by asking players to map the branches of single neurons in a three-dimensional puzzle game.

2.2 Gamification

Gamification is a term "*mired in diverse meanings and contradictory uses*" [Seaborn and Fels 2015]. However a frequently used definition stems from Deterding et. al': "*[u]sing game design elements in non-gaming contexts*" [Deterding et al. 2011b]. Other definitions refer to the need for gamification to invoke similar psychological experiences that a player would feel when playing a 'traditional' game, or "*a process of enhancing a service with affordances for gameful experiences in order to support user's overall value creation*" [Huotari and Hamari 2012]. Going further, Hamari et al. [2014] conceptualises gamification as having three parts in order to connect the existing literature on the topic: (1) **Motivational affordances** refer to what specific features have been implemented to make the system a game and increase motivation; (2) **Psychological outcomes** are the resulting variables that have been measured relating to motivation, attitude and enjoyment, whereas (3) **behavioural outcomes** are those most typically reported in studies and relate to the resulting actions of the player [Hamari et al. 2014].

2.2.1 Game Elements. As a first step towards understanding gamficiation better, Deterding et al. created a taxonomy of elements that are used in projects to frame a task as a game. These game design elements may fall into the following levels of abstraction: (1) game interface design patterns, (2) game design patterns and mechanics, (3) game design principles and heuristics, (4) game models and (5) game design methods [Deterding et al. 2011a]. Building on this, a number of specific game elements are then outlined in Seaborn and Fels [2015]: points, badges, leaderboards, progression, status, levels, rewards and roles. Therefore the popular definition of Deterding above relates to the use of these features in non-gaming contexts. For this paper, we focus on their use more specifically in VCS projects and we investigate the extent to which they have been successful.

Points are referred to as a numerical metric of progress, representing the user's current score. **Badges** or trophies are visual icons awarded for completing some form of achievement, whether this is from completing a specific task or accumulating a certain number of points. Their use has been studied in many situations to increase engagement with Web-based sites (e.g. Easley and Ghosh [2016]). In VCS projects, badges may be assigned for performance-contingent reasons as in EyeWire, where high accuracy and point scores are rewarded, or simply for exceeding a set threshold of contributions, as in the Notes from Nature project [Reeves et al. 2017]. Players are ranked by their score on **leaderboards**, and each individual's **progression** through the game is determined by the various milestones or key tasks they have achieved, such as completing certain levels or alternatively "levelling up". Depending on their current progression, a player may receive a particular title or **'status'**, which may increase as the player traverses increasingly challenging **levels** or stages in the game. Players are motivated to continue in the game in the hope of receiving **rewards**, which may be real-world items or items to use within the game itself. Finally, a player may take on one or more **roles** within the game, depending on the type of character they play as [Seaborn and Fels 2015]. Alternatively, players may be assigned roles based on previous behaviour and achievements, allowing access to new tools or opportunities for contribution [Reeves et al. 2017]. Taking these game elements as a starting point, we examine their features according to a number of theretical frameworks on gamification in Table 1.

Element (mechanic)	**Element (dynamic)**	**Motives**	**SDT concept**
Badges	Collection	Achievement	Competence
Points	Collection	Achievement	Competence
Progression	Acquisition of status	Social recognition	Competence / Autonomy
Status	Acquisition of status	Social recognition	Competence
Levels	Acquisition of status	Social recognition	Competence
Leaderboards	Competition	Social recognition	Competence
Rewards	Collection	Achievement	Competence
Roles	Collaboration	Social exchange	Competence / Relatedness

Table 1. Game elements listed by Seaborn and Fels [2015], associated with dynamics and motives from Blohm and Leimeister [2013] and the self-determination theory concepts outlines in Aparicio et al. [2012] according to common aims and impacts on participants.

In their review of gamified e-participation tools, Thiel presents an overview of different game elements found throughout their literature search [Thiel 2016; Thiel et al. 2016], made up of: achievement (e.g. badges), points, status (or levels), spaces for expression and creativity, feedback (such as notifications), personalisation (through profiles, for example), challenge (missions, quests, etc.), competition, and time constraints. Thiel's overview shows similarities to the list of elements presented by Seaborn and Fels [Seaborn and Fels 2015] which contain many of these features.

There are further characteristics of gamified projects that can be found in the literature but which do not necessarily constitute 'game elements'—they merely relate to design decisions of the project around the game, without actually impacting the core mechanics of the game itself. **Narrative** is a key element in many games; it is the series or sequence of events that help to guide the player through the game's story [Rudrum 2005]. The ability to **skip** is another game feature that allows players to move on to another task with completing the one they're currently assigned to. Skipping is assumed to help to keep players motivated if they encounter a particularly challenging or confusing stage or level in the game. The intended usage of the skip functionality suggests that it provides players with a sense of autonomy as described in SDT [Deterding 2011]. Additionally, numerous games will offer a **tutorial** or practise process that aims to get players started with the tasks involved and is necessary in more complicated games - for example, by completing a specially designed set of puzzles or tasks before proceeding to the true scientific task [Curtis 2015]. Those without a tutorial can be labelled as self-explanatory, indicating that the task is simple enough that players can generally begin

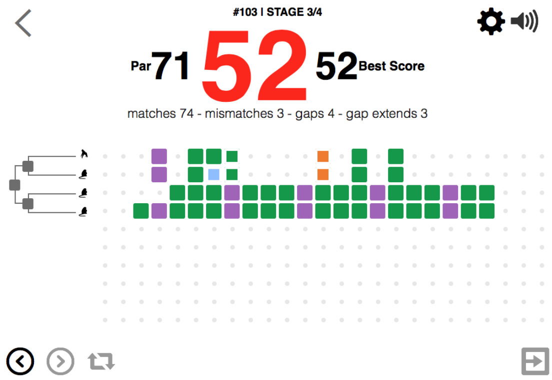


Fig. 2. The Phylo game interface: http://phylo.cs.mcgill.ca.

playing without any formal introduction to the task. Many games also offer the opportunity for **learning** and this is common in VCS projects as the games are based around scientific research, and the process of educating citizens is motivating for both the scientists and the volunteers themselves [Rotman et al. 2012].

2.2.2 Gamification for Engagement. This section details the use of gamification in a range of contexts - for a discussion of gamification in VCS, please see the following section. As discussed in Section 1, there is a debate about whether offering extrinsic rewards is a good strategy for increasing motivation. Deci et al. warn that in many situations these could negatively effect people's intrinsic motivations [Deci et al. 1999], whereas others argue that a game itself is intrinsically motivating [Deterding 2011]. Motivated by this debate, Hamari et al. analysed 24 empirical studies on gamification and found that, generally, the technique does work, albeit with some caveats that were often to do with positive effects only being found for a particular type of use or user [Hamari et al. 2014]. In addition, a review of studies about gamification's use in crowdsourcing found that its use typically leads to an increase in engagement, output quality or other related positive effects [Morschheuser et al. 2016], while an extended review supported these findings to show that it increases participation and quality of crowdsourced work [Morschheuser et al. 2017]. There is an additional consideration that context may well be key, and that certain systems that rely primarily on rational behaviour are not suited to being gamified [Hamari 2013; Hamari et al. 2014]. Finally, the results in Hamari et al. [2014] show that people interact with games and gamified systems in different ways, and therefore they are likely also to have different experiences with a system that incorporates game elements.

Gamification appears to be popular for younger participants in crowdsourcing, but less so for other ages groups [Baruch et al. 2016]. Instead, some participants are drawn to these systems through **altruism**, and their engagement is maintained through egoism and collectivism, rather than competing to move up a leaderboard [Baruch et al. 2016]. Other projects may focus on **sociability**, or as described by George Simmel, the "*sheer pleasure of the company of others*" [Ducheneaut et al. 2007]. This interaction may be afforded through features such as forums and chat boxes which facilitate discussion among participants. Similarly, collaboration is a way of framing the game that requires people to work together to complete a task. These are alternative ways that a game may be framed rather than the standard **competition**-based mechanisms.

2.2.3 Gamification in Virtual Citizen Science. Gamification has emerged on to the VCS scene partly due to cheaper technology, but also thanks to personal data tracking and the familiarity people now have with 'games' as a medium [Deterding 2012; Seaborn and Fels 2015]. There is already a growing body of literature discussing the merits and advantages of gamification in individual citizen science projects, with findings suggesting increased success in attracting Millennials, and end-results such as increased community awareness [Bowser et al. 2013]. However, despite the generally well-received nature of gamification, studies have suggested that certain aspects of it may turn away certain participants—for example in the Old Weather project there appears to be a disparity between those participants who are driven to compete thanks to the use of certain mechanisms, and those more casual users who ignore them or, in some cases, decided to completely withdraw from the project [Eveleigh et al. 2013]. This finding is supported by research into gamification in crowdsourcing which suggests that it could detract from the experience for certain demographics [Baruch et al. 2016]. Where the literature is lacking is in broad assessments of the insights from existing projects that collectively could shed light on whether gamification is typically an effective way to increase volunteer performance with VCS. Metrics have been elicited, split around two main elements (contribution to science and public engagement), for measuring the success of online citizen science projects in general, but these have yet to be used for measuring the effects of gamification [Graham et al. 2015]. There is a distinction within the literature between gamification and gamified projects - i.e., those which take scientific tasks and apply game elements - and Games With a Purpose or serious games - i.e., highly game-like activities such as puzzles which are applied in scientific research.

3 LITERATURE REVIEW

This section details the results of a literature review conducted across four databases. We begin by detailing the methodology used to find and sample publications. In terms of results, our findings discuss empirical evidence detailing the impact of gamification on volunteer motivation and distinctions with less gamified contexts, impacts on output data and the impact of specific features.

We group findings from the relevant research papers below, categorising the outcomes under a number of themes relating to how games can be employed within VCS. These themes were determined through discussion between the authors, based on commonalities in the abstracts, findings and keywords of the research papers. We summarise the findings in Table 7, using the conceptualisation of gamification outcomes presented in Hamari et al. [2014].

3.1 Methodology

We reviewed the existing literature for papers detailing the experiences of utilising game elements within VCS, focusing on those which offer evidence into the impact that this approach has made. While literature surveys have previously been carried out on empirical studies into gamification in general (see [Hamari et al. 2014]), showing that it does tend to produce positive effects, there has yet to be any specific look at what the effects have been in the context of VCS which is important as effects appear to be context-sensitive.

We selected four databases: the ACM Digital Library, JSTOR, SCOPUS and the Web of Science service. Two search procedures were conducted - one search for "citizen science" and "gamifcation" and a second, separate search for "games with a purpose". Each search was carried out to cover all publications up to and including those published within 2017. More specific information on the search terms used and the number of results can be seen in table 2. A total of 517 results were collected across the four databases, of which 344 results were unique.

As it was our intention to find specific evidence for the impact of gamification features, we introduced a set of selection criteria aimed at filtering irrelevant and theoretical sources. The results of each search were analysed by

Table 2. Databases and search terms used and results generated.

Repository	Search terms	Results
ACM Digital Library	recordAbstract:(+"citizen science", "gamification")	13
	recordAbstract:(+"games with a purpose")	62
JSTOR	("citizen science" AND ("gamification")	1
	("games with a purpose")	12
SCOPUS	(TITLE-ABS-KEY("citizen science)) AND (TITLE-ABS-KEY("gamification"))	35
	TITLE-ABS-KEY("games with a purpose")	302
Web of Science	TOPIC:("citizen science") AND TOPIC:("gamification")	17
	TOPIC("games with a purpose")	87
Total	N/A	517

Table 3. Number of Papers by Category

Count	Description
189	Duplicated or highly similar publications
68	Not empirical studies
47	Not peer reviewed academic papers
3	Not related to gamification
82	Not related to citizen science
24	Results or description of systems
44	Irrelevant topics

two separate researchers and overseen by a third researcher who identified and resolved any disagreements between the results. Papers were initially assessed for relevance based on title, keywords and abstracts and duplicated results, as well as those with no relevance to gamification or citizen science were removed from the sample. The remaining publications were analysed at the full text level to ensure the presence of empirical evidence and relevance to the topic of the effectiveness and impact of gamification features in citizen science. The decision was made at this stage to remove those publications describing the results of systems, due to the lack of specific evidence for the effectiveness of particular features and methods of implementation. Table 3 details the number of papers removed, as well as the reason for these removals. A final total of 54 papers was selected for review, which can be seen in table 7.

3.2 Competition Versus Collaboration

In the Old Weather project, surveys and interviews were carried out with participants to discuss the appeal of the gamified approach. The results revealed a conflict between those who found the mechanisms motivating, and those who were turned away by the same features [Eveleigh et al. 2013]. Point scores were suggested to be demotivating to medium and lower scoring participants and the title, while attempting to achieve a higher title by proceeding through the collection-based leaderboard system was stressful, leading to participants to switch to less competitive collections. Moreover the existence of points and reward systems such as badges and titles encouraged some participants to leave the project, as they found such elements to be trivial – a characteristic that was at odds with the serious nature of the research involved. Evaluating the Happy Match project, Crowston and Prestopnik [2013] found that there was a slight indication that participants would be motivated by additional competitive elements, again suggesting that this is a feature that is not universally welcomed or rejected. Other studies have shown that player performance may be increased by showing the progress of a 'virtual peer' whose progress offers a means of comparison for the participant [Laut et al. 2016]. It has been proposed that more granular and personal scores and targets are required to ensure participants feel valued throughout, with an emphasis on quality rather quantity [Eveleigh et al. 2013]. Additionally, team-based challenges or prizes could help to ensure that some volunteers' interest is maintained, negating the problem of some users feeling they were already too far down the rankings [Eveleigh et al. 2013].

Alternatively, [Curtis 2015] found that 'friendly competition' was not a driver for players to participate in the Foldit game; contributing to science and an intellectual challenge were instead much bigger motivations. Curtis [2015] also suggests that game elements offering interaction and collaboration may be more motivational than competitive elements, again demonstrating the different attitudes towards these types of games. Baruch et al. [2016] states that while gamification is popular among younger participants of the crowdsourcing platform Tomnod, it detracts from the experience of others, and cooperation is generally more desirable than competition for all users. This may be related to different types of contributor: Eveleigh et al. [2014] discuss the difference between high contributors who are deeply engaged with the competitive elements of Old Weather, and low contributors who tend to only try out projects for short periods. Eveleigh et al. [2014] emphasise the need to encourage "long-term dabbling", suggesting that breaking tasks into smaller-scale problems can help appeal to the majority of participants' short attention spans.

3.3 Maintaining Engagement

Opportunities to build social relationships with fellow players through game activity have been shown to encourage long-term, lasting participation. Curtis [2015] details the formation of friendships between FoldIt players and the impact that these friendships have on encouraging players to continue to play the game. Similarly, these team- and community-based relationships were also observed to be essential to overcome difficult puzzles and to learn to contribute to a relatively difficult scientific problem [Curtis 2015]. In particular, team formation was noted to enable newer players to overcome the demanding and lengthy tutorial process, thereby enabling and encouraging continued participation by preventing new players from becoming frustrated [Curtis 2015]. Similar findings were reported by Iacovides et al. [2013] from exploratory interviews with FoldIt and EyeWire players. Players reported that the chance to join and contribute as teams was essential for their continued participation in FoldIt [Iacovides et al. 2013]. While EyeWire does not allow explicit team formation, a survey of player motivations for chat participation found that participants nevertheless found ways to collaborate and help one another — particularly when helping newer players learn difficult tasks, as in FoldIt [Tinati et al. 2015].

The impact of extrinsic reward mechanisms was more mixed. In interviews with Iacovides et al. [2013], players described the urge to extend play sessions to maximise progress, as indicated by point and leaderboard mechanics. However, this was suggested to result from a desire to be part of assisting scientists in completing research that would otherwise be impractical — rather than any interest in point or leaderboard rankings. This is echoed by findings from the Quantum Moves platform, where the presence of leaderboards was demonstrated to have no statistically significant impact on players' likelihood to return to the game [Pedersen et al. 2017]. Conversely, Prestopnik et al. [2014] et al suggest that points-based gaming is effective in encouraging long-term play. Prestopnik et al. [2014] noted that points-based games were associated with higher volunteer retention rates than progression and narrative-based gaming. A follow-up study by the authors suggested that this results from players exhausting progression opportunities [Prestopnik et al. 2017]. While players can eventually exhaust progression and narrative mechanisms, the opportunity to earn points persists as long as players contribute to games. The opportunity to build scripts to outperform fellow players and maximise point gains has also been stated by some FoldIt players to be key to their continued interest in the game [Ponti et al. 2015].

Once a player has been attracted to the game, there are differences in the way that the game itself is implemented that can affect participation. Story-based games have been shown to be preferable compared to points-based games [Prestopnik and Tang 2015]. Participants were motivated to continue playing as activities were 'story-focused', suggesting that these types of game may be particularly effective for VCS projects, especially if the player is not interested in the scientific outputs [Prestopnik and Tang 2015]. The authors suggest that this type of game could help to balance the contributions that are typically skewed towards a few members of the community—a community that is typically made up of people interested in the science rather than just playing a game [Prestopnik and Tang 2015].

Social Factors and Cooperation Iacovides et al. [2013] found that the provision of discussion features such as chat and forum platforms was important for developing a sense of community. These opportunities for team interaction and for team-play were important motivators for continued player engagement. Further, Eveleigh et al. [2013] suggest that the implementation of team-based activities and rewards could be used to reduce the potentially demotivating impact that competition has on underperforming individuals. This is echoed by Baruch et al. [2016] who note that users of the crowdsourcing platform Tomnod responded more favourably to cooperative elements than to competitive elements.

Competition The inclusion of competitive elements has been found to lead to increased levels of contribution, but lower levels of effort in tagging tasks - while players generate more tags, the time taken for each tag and the quality of tags is lower in competitive games [Siu et al. 2014]. The impact of competition is also highly dependent on the performance of other players. While players will display increased participation in the face of competition, participation will decrease if fellow players are deemed to contribute too much or too little. Experimenting with a virtual peer, Laut et al. [2016] found that players adjust their behaviour to match peers, reducing contributions when peers are less active and increasing contributions as peer's activity increases. This effect is only seen, however, where such levels are seen as achievable. Opportunities for competition can be particularly demotivating to casual participants and newcomers, who find themselves outperformed by more active and longer-term players, leading less active players to leave projects [Eveleigh et al. 2013].

Feedback Evidence of progress was identified by participants of FoldIt and EyeWire as a rewarding element of games that positively reinforces player participation, leading to increased activity [Iacovides et al. 2013]. Feedback features have also been found to assist players in adapting the way they contribute. Fylo contributors found that game elements provided additional information which was useful for informing their behaviours [Forde et al. 2016]. Nevertheless, this had no significant impact on the players' intrinsic motivations, nor their competency, suggesting the impact of such features is minimal.

3.4 Participant Motivation

The appeal of participating in science was strongly emphasised in the literature, particularly by Iacovides et al. [2013]. In the EyeWire platform, Tinati et al. [2016] surveyed participant motivations in the citizen science game Eyewire, finding that the most significant factor influencing player participation was the desire to contribute to scientific research. This is similar to other VCS research that has shown that a desire to contribute to scientific research is a strong motivation for participation [Raddick et al. 2013]. Jennett et al. [2016] note that the factors governing initial attraction to citizen science almost exclusively pertain to scientific research - interests in science, curiosity about research and a desire to contribute to research. Nevertheless, continued interest in projects is a significant factor influencing long-term player engagement [Jennett et al. 2016] and our findings suggest that it is in this area that games have a significant influence.

Perspectives on the motivational value of game elements were mixed. Forde et al. [2016] suggest that the addition of gamified elements has no significant impact on intrinsic motivations. Similarly, from Iacovides et al. [2013] in interviews with participants in EyeWire and FoldIt, it was found that game elements did not necessarily attract players to the project—their intrinsic interest in the science was enough—but such elements had some motivational value in providing recognition that participants were contributing. Similarly, Curtis [2015] found that FoldIt players were predominatly motivated by altruistic desires to contribute to science, rather than extrinsic mechanisms such as points and competitive elements.

Reward-based game elements such as points have been noted to impact negatively on those participants who show strong intrinsic motivations [Prestopnik and Tang 2015]. Even among those players who are more extrinsically motivated, Siu and Riedl [2016] describe that gaming alone was insufficient to overcome the boredom associated with repetative human computation tasks, particularly where players derived fun from collecting extrinsic factors such as game unlockables. On the other hand, Bowser et al. [2013] suggest that reward-based game elements such as badges are associated with player interest and enjoyment and that as a result, these features do in fact increase the motivation to participate.

Further, in spite of the importance of altruistic and intrinsic interests in science suggested above, Greenhill et al. [2016] describe highly gamified features with no scientific purpose. The Zooniverse's *Vorwerp Pong* drew significant interest from project participants, despite lacking any scientific purpose. Greenhill et al. [2016] suggest that both intrinsic factors such as a desire to have fun and extrinsic factors such as points, leaderboard rankings and rewards are key to driving participation in gamified citizen science.

3.5 Output Data Quality

In some cases, project teams have suggested that introducing gamification elements has been detrimental to the quality or accuracy of the data produced[3]. A comparison of GWAPs with microtask crowdsourcing platforms conducted by Thaler et al. [2012] found that the proportion of correct responses submitted within the GWAP was 15% lower than within the crowdsourcing platform (74.87% compared with 89.75%). However, a similar study carried out by Sabou et al. [2013] found the performance of crowdsourcing platform participants to be highly interface-reliant, with game players matching or even exceeding the performance of microtask crowdsourcing. Within human computation, Chung et al. [2017] have also observed differences in player performance based on the visual load of tasks. By providing participants with more information and more on-screen subjects for analysis, players were able to better formulate gameplay strategies, with the mean performance of players increasing by 80%.

Aiming to address some of these concerns, Prestopnik et al. [2014] compared the quality of data produced by one gamified project with one that was presented as a game that required users to complete a task in order to progress (note the differentiation between a project that includes '*gamified*' elements, and one which presents itself as a game, which we return to later). While the game-based project did introduce cheating as a strategy among players, overall (excluding the cheating behaviour) there was no significant difference in performance [Prestopnik et al. 2014]. Similar findings have been made in other domains, around adding competitive elements in crowdsourcing to improve volunteer output

[3]For example see http://www.cancerresearchuk.org/support-us/citizen-science/the-projects#citizenscience1

without harming the quality of the data they produce [Rokicki et al. 2016]. Taking a somewhat different approach, the effects of collaborative and competitive game elements have been studied, based on how they affect player engagement and accuracy in games with a purpose (GWAPs). Results suggested that as well as improving engagement, competition produced results that were as accurate as collaboration [Siu et al. 2014]. Of note was the finding that when combined—when a player is presented with both competitive and collaborative scoring options—performance levels decreased [Siu et al. 2014]. Similarly, Goh et al. [2010] demonstrate that competitive models outperform collaborative models in image matching task, due to players' rewards and progress being more dependent on the quality of submissions. This is not the case for all activities, however. Framing image matching tasks in combined collaborative, competitive game-based activities has been shown to increase the quality of labels generated by players [Huang and Fu 2012].

Nevertheless, concerns remain about the impact of competition on data quality. A number of studies surveyed associated competitive elements with reductions in the time spent by volunteers on each contribution, raising the possibility that participants could rush, introducing errors to data [Mekler et al. 2013; Siu et al. 2014]. Furthermore, players showed an unwillingness to provide some forms of data, regardless of the rewards offered by games, leading to incomplete data submissions [Kapenekakis and Chorianopoulos 2017]. While game features generally had little effect when attempting to increase the accuracy and quality of submissions, studies suggest that players are influenced by the way in which tasks are framed. Presenting tasks in a manner which emphasises their scientific value and players' altruistic motivations encourages players to spend greater time and effort on each subnmission — leading to an increase in the quality of submissions without reducing engagement [Mekler et al. 2013].

4 CITIZEN SCIENCE PROJECT SURVEY

4.1 Methodology

In order to obtain a more complete picture about the uptake and use of these elements, we surveyed and observed a number of projects, and quantified how frequently each type of game feature was used. To focus only on "virtual citizen science projects", we used the typology presented by Wiggins and Crowston [Wiggins and Crowston 2011] that suggests such a project should require no physical elements whatsoever (for example, paper-based data collection). Therefore we targeted only those projects in which all stages were carried out via digital means.

We subsequently consulted the human-curated list of citizen science projects on Wikipedia, and the same list for human-based computation games, which we then manually filtered to obtain a sub-list of active VCS projects. The lists were retrieved from Wikipedia in November 2015. Projects were removed if they involved offline activity (97), the project website was closed or had failed at the time of study (4), the website for the project had ceased (5), there was no English language version available (2), or in the case of the human computation projects there was no proof that it was a citizen science project (1). In addition, there were a number of projects that appeared in both lists, so duplicates were removed (4).

Across both lists we found a high number of incidences of projects drawn from the Zooniverse platform. These projects were generally very similar, with few (if any) gamified elements and represented the majority of the sampled lists. In these cases, we made the decision that where there were numerous projects from the same platform, each of which typically used a similar set of features and shared many characteristics, we took a subset of these projects. Such restrictions were particularly necessary given the possibility that just the one, predominantly non-gamified platform would have a much more significant impact on findings. We obtained a final list of 31 VCS projects that we then assessed in order to establish the prevalence of different game elements. The final selection (see Appendix 8) is not intended to

Game Elements	Game Characteristics
Points	Narrative
Badges	Altruistic framing
Leaderboards	Sociability
Progression	Tutorial
Status	Collaboration
Rewards	Competition
Roles	Learning
-	Option to skip

Table 4. Game elements and characteristics surveyed in each project

Game Element	'Game' Projects (N=14)	Non-game Projects (N=17
Points	13	4
Badges	5	4
Leaderboards	11	6
Progression	7	2
Status	5	1
Rewards	4	4
Roles	1	3
Narrative	5	1
Altruism	11	15
Sociality	6	9
Tutorial	7	12
Collaboration	4	8
Competition	12	5
Learning	3	13
Skip	6	6

Table 5. Survey results for projects describing themselves as games and projects not described as games.

be an exhaustive list of all VCS projects (particularly because we filtered many similar Zooniverse projects); rather we were interested in finding those projects with sustained activity (indicating sufficient popularity to avoid closing) that could be used to assess design decisions.

For each project, we assessed whether or not it included certain game elements or characteristics, specifically: points, badges, leaderboards, progression, status, rewards, roles, narrative, altruistic framing, sociability, tutorial, collaboration, competition, learning, skip (also depicted in Table 4).

4.2 Results

The previous section shows that the literature has a generally positive perspective on a number of game elements in certain projects. Our survey of projects revealed the extent to which various elements were employed by the 31 projects in our sample, with the results depicted in Table 5.

4.3 Comparison with Literature Findings

Overall, our results demonstrate that the game mechanics of citizen science games predominantly rely on providing extrinsic rewards for participation — rewards such as points and leaderboard rankings. As points are probably the most basic form of tracking a user's progress through the game, this is perhaps not surprising; badges, progression, status, rewards and roles are all features that would in most instances need to build on the implementation of points in the first place. Nevertheless, although the findings of our literature review suggest that cooperative elements motivate greater participation than competitive elements, almost all of the 14 game projects leveraged competition as a game mechanic. Similarly, in spite of the effectiveness of narrative and progression based gaming, these features were relatively uncommon and closely linked — progression was exclusively linked to proceeding through narrative arcs. Opportunities for social interaction were present in 50% of games, with few opportunities for group and team forming, contrary to the positive role these features can play in encouraging long-term participation.

While learning was not a motivation associated with gaming, we found a small number of games that offered players the chance to learn more about the underlying science associated with a game. In addition, perhaps contrary to the reliance on extrinsic motivations, the majority of projects framed tasks in terms of altruistic values and the importance of tasks for science. Tutorial systems were present in around half of the games surveyed, to assist players in learning how to contribute. These systems varied in complexity, length and the effort demanded of players — FoldIt featured the longest tutorial process. It should be noted that while tutorials helped players 'learn' to play their respective games, these tutorials were not learning features — none of the tutorials explained the underlying scientific research.

4.4 Distinguishing Games from Non-game Projects

In order to extract information about how the projects were described, compared with what elements they used, we also carried out a similar frequency check on the projects by splitting them on whether or not they described themselves as a 'game'. We searched for each project using Google, along with the word 'game' and for those projects which provided no result (e.g., they listed the project, but with the word 'game' missing), we then looked at the project website for references to it being a game. 14 of the 31 projects were found to describe themselves in some way as a game. In Figure 5 we see that for the game elements from Seaborn and Fels' (points, badges, leaderboards, progression, status, roles), those projects that describe themselves as 'games' do, unsurprisingly, make use of the elements to a greater extent than those which do not. However, for many of the other game characteristics, it is the 'non-games' that more regularly demonstrate their use—exceptions to this are narrative and competition, which are much more popular in 'games'. Interestingly, it is the projects that do not describe themselves as games which focus more on the learning and education aspects of the work and as such suggests that these projects may instead be hoping to engage with volunteers in a different way, reflecting the suggestions from the literature that projects may offer success through different forms of framing (e.g. Curtis [2015]). This echoes the tensions described by Prestopnik and Tang [2015] and Eveleigh et al. [2013], among others, in terms of the impact of framing 'serious' citizen science tasks as games. While games frame their tasks through the awarding of extrinsic factors for participation, non-games instead support intrinsic motivations by reinforcing players' intrinsic interests through learning and feedback [Jennett et al. 2016]. There are, however different degrees of such framing and it should be noted, that not describing the project as a game does not mean that it won't include any game elements at all. We therefore recognise an interesting difference:

(1) 'Games' are those projects that are openly described as being a game, and include the implementation of game elements and characteristics;

(2) 'Gamified projects' do not describe themselves as games, but include some kind of game mechanic;
(3) 'Non-gamified projects' rely only on altruistic motivations and do not seek to improve user engagement through the use of game elements.

5 INTERVIEW FINDINGS

Building on the outputs of the studies discussed so far, this paper aims to provide in-depth qualitative data about the role of game elements in a VCS project, for which interviews were carried out with the EyeWire project team.

5.1 Methodology

Semi-structured interviews were run during February 2017 with members of the design team for the EyeWire project–a gamified citizen science approach to mapping the human brain (see Figure 1). Three interview sessions were carried out in February 2016 by a single researcher, covering six participants who represent the entirety of the full-time staff working for the EyeWire project at the time. The six members of staff were:

(1) Community and Project Manager - Responsible for overseeing the project, this participant was interviewed alone during interview session one.
(2) Designer - Responsible for designing artistic and interface elements, this participant was interviewed alongside the developer during interview session two.
(3) Developer - Responsible for implementing new design elements and resolving system bugs, this participant was interviewed during session two.
(4) Game Master - Responsible for interacting with the community, planning and running community-engagement activities and elicting feedback from the community. Interviewed during session three.
(5) Game Master - Interviewed during session three.
(6) Game Master - Interviewed during session three.

The purpose of the interviews was to understand the thought processes and design decisions taken by the citizen science designers relating to gamification elements, in order to provide further details about why these elements were introduced, how they are assessed by the project team (in terms of their performance relevant to the aim of the project), and whether or not it is felt that they are ultimately successful. We followed a semi-structured schedule in order to ensure that a key set of questions were asked and answered [4], but wanted to allow the participants to deviate where necessary so that we could gain insights into the entire process of design and the reasons behind the introduction and use of gamification. We used thematic coding to analyse the interview transcripts [Pope et al. 2000], eliciting the concepts and themes discussed by each participant that help shed light on the perceived success of the game elements.

Interviews were transcribed by a single researcher, before being checked and corrected by a second researcher. These transcripts were initially coded by a single coder using NVivo, with a qualitative codebook used to allow flexibility as new themes were discovered within the interview transcripts. The process used was iterative - An initial set of codes was devised in anticipation of opinions and thoughts about the use of game elements and their effect on the citizen science project as a whole. A second thematic coding analysis was used based on this initial set of codes to extract the underlying themes from the participants' responses. The transcripts were then coded independently by a second researcher. Differences between the two coding schemas were analysed by a third researcher, responsible for resolving these differences. The remaining findings within this section are grouped according to the emergent themes identified

[4] A full list of these questions can be seen in appendix B

during the thematic analysis process. A summary of the interview findings and the relevant respondents grouped by these themes can be found in Table 6

5.2 Decision Making

Given the limited time and development resources available for the project, a common theme mentioned throughout the interview process was the need to carefully consider which features and changes should be made to the EyeWire project. In spite of the scientific needs of the project, the team predominantly described player-driven changes to gameplay and game features, rather than to the underlying scientific infrastructure of the project.

"We have completely redesigned the interface of Eyewire, we've designed a different onboarding system that's gamified, has narrative, has art that goes with it. We've built a notification system, a pop-up system, a better infrastructure for joining teams. An idea for how players could form their own teams and compete. We have reimagined how you would have an automatic level progression system. Most of the things that we've wanted to build are more on the game mechanics side." (Participant 1)

In truth, the team rarely made use of quantitative data for decision making. This was predominantly reserved for contexts in which development time would be spent on scientific needs, where players would be largely unaware of the changes made.

"We're more influenced by player feedback." (Participant 2)

Usually if we're doing data-driven feature development or feature iteration, it's in the interests of efficiency, or proving that this is a feature that's needed." (Participant 1)

Nevertheless, this reliance on player feedback raises a number of issues. The first of these is that the team found it much easier to obtain feedback from those players who are highly engaged to the point at which they are considered 'top players', whereas there aren't as effective mechanisms for eliciting these views from newer volunteers who perhaps aren't as involved yet in the social side of the project. This has led to a development schedule focusing on maximising engagement from existing users and a resulting trade-off in engagement from newer and less active players.

"We have over time probably grown more towards building more features for power players. And that has had a negative consequence of not dramatically improving our onboarding for new players." (Participant 1)

"[H]ere we're trying our best to sustain the community, give them what they need and keep the project moving forward." (Participant 3)

Moreover, a significant insight raised by the team was the difficulty in ascertaining the accuracy of feedback given by the community. Of particular interest was the anecdotal description of the introduction of the *Scythe Complete* feature, where the team had worked closely with top players to understand their views. Initially the scientists behind the project had hoped that the new feature would increase the accuracy of results and players had expressed their interest in being able to further assist with scientific research. Ultimately, however, the feature went unused as players preferred the better rewarded task interactions, with a greater bearing on their leaderboard ranking and point score.

"[Our mistake] was to listen to what people say, not what they do. The players were enthusiastic about it, they thought it was cool. [But] they like to say that they don't play for points, but they do" (Participant 1)

5.3 Game Elements

Throughout the interviews, our focus was on understanding how and why the participants selected gamification elements to add to the project. The main elements whose use was described by the interviewees were points, leaderboards and badges. However additional game features, including chat, competitions and accuracy bars (for showing performance) were also discussed.

Points, generally a crucial aspect of games and required for other game features, were described in all interviews as the fundamental element that drives the game within EyeWire.

"The most important? Probably, you know, points is up there. . . . because you know, your core task in Eyewire is to map a branch of a neuron from one side of the cube to the other and the reward that you get for doing that cube is you get points." (Participant 1)

"Anything that contributes towards your points score is definitely more incentivised than things that don't." (Participant 4)

However, whether or not the use of points is seen as an effective means of increasing engagement is uncertain. While the respondents were positive about their influence on driving a certain portion of the player-base to engage, it was not clear how effective they were for all players. In particular, it should be noted that EyeWire is very much a point-based game. Points were a fundamental part of EyeWire from the very earliest phases of the game and at no stage had the team conducted comparative testing on the effectiveness of other, less extrinsic elements. Rather than drawing on any specific evidence, the team's opinion on the use of points predominantly stemmed from their views of participant motivations:

"Are they motivated by points? Are they motivated by joy? Are they motivated by the community? I think that we have players that are motivated by all of these categories." (Participant 2)

"I think people like the points score the most, not just because it's on the leaderboard, but because it's kind of the core interaction that they're familiar with, but also you can... It's a value that can accrue the fastest in a sense." (Participant 5)

Overall, the EyeWire team expressed the view that they felt participants in VCS games expect to be sufficiently rewarded for their time and effort. It was this philosophy that drove many of the subsequent decisions to implement and modify features. Elements such as badges were added to reinforce the impact of earlier reward systems, ensuring that players did not exhaust the point systems. The team expressed the concern that accumulating points with no sense of getting closer to any victory would not otherwise provide the motivation to continue, requiring the introduction of further elements to encourage engagement with the project over a longer period of time.

"Any game, if it's a game, generally speaking needs to have a reward system." (Participant 4)

"[P]oints can be... They can seem a little arbitrary and you know, once you have tens of thousands of points, you know what's another one or two hundred that you get from doing a cube? So we introduced badges." (Participant 1)

However, this raises something of a contradiction. A common theme expressed by the team was the limited development time available for the project. Conversely, by the participants' own admission, iterative development was required to regularly produce new badges, or achievements, that players can play towards and remain engaged. Participants suggested that this reliance on reward mechanisms was largely due to the difficulties perceived by the team in modifying the scientific task to allow progression or introduce narratives:

"Some of them, [the ultimate aim] is just ultimately winning. In our case it's points, you get to be above your other peers in showing how successful you are. Because you can't really win Eyewire." (Participant 4)

Nevertheless, the team did describe a small number of progression-based features. In particular, this lack of an opportunity to win led the team to introduce smaller goals and forms of progression — specifically competitions and accuracy bars displaying a percentage accuracy score for each player. While the developers believed that players were motivated a lot by the accuracy bars that EyeWire uses to display the 'quality' of a user's contributions, this feature caused mixed effects as players became more concerned with maintaining a good score on this as opposed to completing certain types of work. The team found it apparent that different levels were required in games to ensure that players of different ability and experience may play the game in a way that is most fulfilling for them, while experiencing a feeling of progress. At the same time, they expressed the belief that players would lose motivation if they felt their own scores would be affected by the ability of others in the game, and so care had to be taken to ensure that similarly-experienced users were grouped together. This was not an easy task given the need to maximise contribution levels from players and to ensure players were not unneccessarily prevented from contributing to the game.

"[T]here is some issues with accuracy being shown, in that certain higher ranked players prefer not to play on easier cells, because they're more likely to be compared with lower ranked players. And that can temporarily affect their accuracy." (Participant 5)

It is notable that many of the features described supplemented, rather than augmented the EyeWire task. The team themselves raised this issue multiple times during the interview process. Questions like what is a game and what do players expect from a game experience were crucial to the team, not only in design but also in the daily running of EyeWire. In particular, the team made the comparison with other citizen science games - notably FoldIt and EteRNA - which they felt had the look and feel of a game, while still retaining the core of citizen science activities. In comparison, despite the perceived value of the game elements introduced to EyeWire, the team nonetheless remained concerned that there were issues surrounding the tasks assigned to players — specifically, that it remained relatively 'dry' and work-like.

"We call it a game, but is it a game?... I think there are some games, ETERNA, for instance, that feel like a game. It literally feels like a great game on the internet, it's engaging, the graphics are cool, the interaction is really fun..." (Participant 2)

Arguably one of the most successful features as perceived by the team was ultimately not a game feature at all, but rather one introduced to assist players in helping one another contribute to the game. The team's perception of player use of the chat feature was highly positive and they noted that it was one of the key elements in overcoming the work-like feel of EyeWire. In fact, the team's perception of chat and its importance for players was so great that they had introduced a *trivia* competition class solely to encourage chat participation — even though the competition itself yields no scientific data.

"I'd have to say chat, because with a small community they get to know each other and it kind of makes up for maybe the dryness of the default work that we're having them do" (Participant 3)

"We also have our trivia, which happens during those competitions that has absolutely nothing to do with game play itself... It gets people excited to talk to each other" (Participant 5)

5.4 Game Evolution

We were particularly interested in eliciting from the team how the game has changed over time and whether there were any procedures that were followed before deciding to introduce a new game element or feature. We have already covered the introduction of new sets of badges to increase the motivation for players who have already collected all of those available, but there were other changes that were discussed related to this topic.

The initial version of EyeWire was relatively rudimentary with few features. Players could earn points and leaderboard rankings, but these would reset after each day. Ultimately, this was the extent of the early EyeWire game interface. Changes were predominantly motivated by a sharp reduction in player numbers:

We launched Eyewire, there was a huge bump in traffic and then after like, a month it just dwindled back down and we were like 'crap, what's going on'?" (Participant 1)

In response to this reduction, the team opted to focus on introducing competitive elements to the game. The reasoning behind this introduction was twofold: firstly, the team felt that competitions were integral elements of any game and secondly, that players were eager to show that they were the best.

"[We thought] 'It's probably because we don't have any competitions!' These people expect a game. Let's give them a game" (Participant 1)

"It was kind of like an ad hoc community thing where one players was like 'I'm the fastest neuron tracer in the world'. We were like 'Oh yeah, want to back it up? Let's do a competition!'" (Participant 1)

Nevertheless, while the team believed in the effectiveness of these competitions, participant 5 explicitly stated that no effort had been made to monitor competition sign-ups. Rather, it appears that from the team's perspective, the introduction of competitions had the desired effect of reigniting interest in EyeWire. Rather than continuing and exploring other features which may have been less successful, the team instead opted to strengthen these competitive elements, adding additional infrastructure and competition models.

In addition to competitive elements and similar to recommendations made within the literature, the team implemented progression opportunities and unlockable features. Players who demonstrate high levels of accuracy are able to access new features and receive indicators of their status to display to other players. This introduction brought in the use of Seaborn and Fels' 'Roles' element, which gives players another motivation to try and progress in terms of their class [Seaborn and Fels 2015]. These roles, however, were predominantly introduced to reduce the workload assigned to the team.

"As the player base proved that they were better, we adapted the game system to give them more power and we give them more power and allow them to have more features by giving them a ranked class. So they become mentors or moderators which is mentors helping other players, moderators policing chat." (Participant 1)

"Because [when] there was not a good system for giving more power to the players who were actually good in Eyewire, it meant that the burden was on us to make sure to be fixing community errors." (Participant 1)

Perhaps surprisingly, the introduction of features to the platform did not appear to follow any exact strategy. Rather than drawing on the experiences of other games or recommendations from the literature, the team instead drew predominantly on their own observations of the community's needs and specific feedback from top players.

"[we ask] what does the community need, what are they struggling with, how can we make their lives either more efficient or more enjoyable? (Participant 2)

There did not appear to be much influence from other citizen science games in terms of evidence of what game elements are effective that led to the introduction of features in EyeWire. However it was suggested that generally the team has an awareness of what other games (not necessarily citizen science games) have and are doing:

"Thinking about what elements other games have, I'm sure that plays into what I think would be a good feature for the game." (Participant 6)

This raises the question of where the team got the idea for specific new features - for example, the reasoning behind the early decision to focus on competitions. Far from being an evidence-based process which is carefully researched and implemented, the team suggested that these decisions were predominantly 'gut' decisions, based on the participants' own views and opinions. Our findings suggest this was for two key reasons - firstly because the team had negative perceptions about the genre of VCS games and secondly because of a lack of resources.

"But I would say, fundamentally, it's more due to what... Like, people here at headquarters think may sense for the game and not necessarily looking at say 'what's Foldit doing right now?'" (Participant 5)

On the one hand, this approach - according to the team - has been relatively successful in maintaining the engagement of these top participants, there has been less engagement from newer players and those displaying average levels of activity. Furthermore, the team described struggling with prioritising features, noting a large backlog of new elements intended to improve the game experience that had yet to be introduced.

5.5 Monitoring Impact

We also investigated how the team track the effects of changes and assess whether the introduction of new game features was positive or negative. Generally, the standard player engagement statistics are examined on an ad-hoc basis to see if there were changes in player activity, and whether these were as a result of a particular change:

"[I]f we know when a certain feature was released, or whatever, we can be like 'oh that spike there was when we released this' or 'oh, we released this then, but I don't actually see much of an impact'." (Participant 4)

"We look over time at our cell completion rate and it's going up and that's good. We don't spend a huge amount of time devotedly scouring through stats and looking at engagement, churn rate and things. We've gone through periods where we've looked a lot at the stats, but the kind of takeaway has been - We know a lot of this stuff, we know the things we need to like improve a lot of these numbers. The problem is we don't have enough development time. So, we've just focused on other things." (Interview 1)

Indeed the qualitative insights are almost more important for the team in order to understand more the effects of a new feature, rather than the statistics about how much that feature is being used which appear to require too much time to track effectively. This emphasises the importance of the chat functionality on EyeWire, that provides a real-time stream of the players' thoughts on what is good or not.

"I think what is more challenging for us is understanding like the softer points of how the players are using a new feature. How they're going about reaping cubes, how they communicate with one another on a tough cube, how we can give them feedback . . . You know, just those how cases are a lot harder, but they're a lot more valuable to us, like how does a player learn to be good at Eyewire? We think it's by doing a lot of cubes. We think it's by talking to other players." (Participant 1)

Explaining further, we extracted a key insight into the mindset of citizen science players and the importance in this case of the tasks feeling like a game, rather than work:

"[T]hey sort of played around with it for a while. But then after a couple of weeks, maybe a month, they weren't really using it very much and so we just kind of started - I just started doing chat interviews with the players and I was like 'why aren't you guys using this'? And they were like 'Oh, well, we just - We dropped down on the leaderboard'. 'We're - What's the point of doing this? It just feels like work'. That was their feedback. It felt like work. And they didn't want the game to feel like work. And I think they felt like work because there was no reward." (Participant 1)

It was in response to these reports from players that the team once again turned to reward mechanisms to make tasks more motivating. Yet we note there remain questions surrounding the extent to which rewards alone can make an unrewarding work-like task feel like a game. This reiterates the earlier point made by the team surrounding the expectations of players - players expect a game, but is that what they receive? While the team were quite confident that players did not expect the game to feel like work, they were overall less confident about where the work/game distinction might be found.

5.6 Is Science a Game?

A common theme throughout the interview process was the existence of a number of issues resulting from reconciling the needs of a game with the needs of scientific research. One tension mentioned by the EyeWire development team is the visibility of features that are introduced. The team described issues arising from the needs of the project in terms of science and the extent to which players will be aware of these changes. In particular, there was a need to reach a balance between the needs of the players and the needs of the project scientists:

"[Suggestions for] features we should improve on is sort of from two directions. We either say what does the community need... And then there's the back end side, which is the scientists at Princeton that say we need to get more work done."' (Participant 2)

This became particularly problematic in the case of adjustments to the project that were required to increase the accuracy of submissions.

"the way that we generated the 3D meshes. It's very new geometry. They'd definitely grown attached to the old way that things look... They perceived the defects as details." (Participant 3)

"[New meshes] had fewer discontinuities in them and they were more scientifically accurate, but ironically the community had grown accustomed to the discontinuities in the old meshes." (Participant 2)

While the game itself does not negatively impact the quality of submissions, players resist efforts to improve the suitability of submissions for scientific research. This is surprising given that a desire to contribute to scientific research is a key motivation among EyeWire players, as well as VCS game participants more generally [Tinati et al. 2016].

A further tension arises from the slow rate at which scientific research advances. Players expressed unrealistic expectations of the speed at which research would progress and expressed disappointment in interactions with project scientists:

"Science goes slow. One of the players was like 'so what have you learned since last time'? And I think one of the researchers was like... 'Oh, we haven't even looked at them' That was the truth but it made the players feel like crap." (Participant 1)

As a result of this, the team have made an effort to shield players from interacting with project scientists. This raises questions about the extent to which participation in science is truly motivating for citizen science participants. Rather than altruistic contributions to science research more generally, observations from EyeWire suggest that players are motivated by contributing to specific scientific discoveries and research outcomes.

A related third and more complex issue concerns the extent to which a successful game is a successful scientific methodology. *"There was definitely a push to make Eyewire more efficient than it currently was. Because for a time, when you really did the maths, it was like... The amount of money that was being put into... Like... Paying employees to run the citizen science game itself was more than the money being put into paying tracers at Princeton to just do it themselves and the quality of the data was not necessarily better on our end." (Participant 5)*

"Have people heard of this, are they interested in it? And then whether or not the project is successful in producing something. There's certain kind of degrees of where you find a project successful." (Participant 4)

A repeated concern expressed by all members of the team was the tendency that players showed to request features that were neither suitable nor feasible for a citizen science game.

"[Players] are very vocal and we usually have a list of things we have to do." (Participant 7)

"Our problem is ... based on trying to get outsiders to help offer kind of new game ideas is that we have various... a lot of restrictions on how we need to do things to kind of get accurate scientific results." (Participant 2)

These findings suggest a further potential data quality issue arising from citizen science games. If the EyeWire project is to be financially and scientifically viable, the team must make every effort to maximise the quantity of contributions made to the project. At the same time, to truly maximise contributions, the team face compromises in terms of the features available in projects and the accuracy of the submissions gathered from players. Moreover, balancing the needs of two different communities - players and scientists - is an important, but complex process. Many of the changes

Table 6. Summary of Interview Findings

Theme	Finding	Respondent
Decision Making	Limited time and resources available to VCS games compared with other genres	1, 4, 5, 6
Decision Making	Player feedback, understanding how features perceived more valuable than player statistics	All respondents
Decision Making	Feedback easier to elicit from most active players - leads to development focused on this group	All respondents
Decision Making	Player feedback does not always match reality of player behaviours	1, 2, 3
Game Elements	Players Prioritise tasks with greater rewards/point scores	1, 4, 5, 6
Rewards	Rewards (e.g., points) lose value over time as players accrue greater numbers	1
Rewards	Progression hard to introduce - no final goal to achieve or obstacle to overcome	1, 5, 6
Rewards	Individual reward elements do not make task into game	2
Rewards	Social interaction effective at breaking up monotonous tasks	All respondents
Rewards	In games, less rewarding/fun tasks abandoned, regardless of importance for project	1, 2, 3
Game Evolution	Individual game elements insufficient to motivate long-term engagement	1
Game Evolution	Competition between players promotes long-term engagement	1, 4, 5, 6
Game Evolution	Players can moderate game and chat activity effectively	1, 5, 6
Game Evolution	EyeWire evolution based on gut instinct, player feedback - not observation of other games	1, 3, 4, 5, 6
Game Evolution	Limited resources lead to slower rate of feature introduction	All respondents
Monitoring Impact	Interaction with/observation of players sufficient to understand impact of features	1, 3, 4, 5, 6
Science vs Game	Projects must balance science and scientist needs with player needs	All respondents
Science vs Game	Introduction of improved features can be demotivating	2, 3
Science vs Game	Slow progress of science negatively impacts on volunteer intrinsic motivations	1
Science vs Game	Fun games not necessarily efficient or cost-effective	1, 4, 5, 6
Science vs Game	Scientific needs limit features which can be added to games	All respondents

introduced by the team were ultimately one-sided, addressing the needs of players with little impact on project outputs or alternatively ensuring the accuracy of submissions while impacting on player engagement.

In the case of EyeWire, at least, while the scientific activity can be made into an engaging and challenging game, our findings suggest that contrary to motivational findings suggested by the literature, it is the game elements - points, rankings and progression - that attract and maintain the interest of players.

6 DISCUSSION

We have investigated the use, reporting of and effectiveness of gamification elements in virtual citizen science. Within this paper, we have provided three contributions: (1) a review and summary of existing literature into gamification effectiveness for VCS, (2) an overview of what game elements are used in VCS projects, and (3) findings from interviews carried out with the EyeWire project team to understand more about how these elements are introduced to the game, tracked and assessed to determine their performance.

6.1 Gamification, Engagement and Motivation

Overall, our findings suggest that gamification and the use of game elements *supports* rather than replaces other motivational factors which are common in less gamified projects. Previous research has identified altruistic motivations as important factors in encouraging players to return to projects and shifting from short-term to long-term contribution patterns [Crowston and Fagnot 2008; Rotman et al. 2012]. While the literature review process provided some evidence that game elements have a similar effect, this was not supported by all sources and moreover, the EyeWire team suggest that extrinsic reward factors such as points can be demotivating in the long-term as players gain increasingly more and such rewards lose context. In fact, experiences that evoke negative emotions - emotions such as distress or sadness - have been suggested to nevertheless motivate player participation in games, if the players feel that they can make a difference through their participation [Iacovides and Cox 2015]. This suggests that altruistic motives are more significant than extrinsic factors and 'fun' for long-term participation.

Furthermore, the literature demonstrates that it is intrinsic factors - specifically interest in the underlying research topic, field or in science itself - that is the single biggest driver of participation in citizen science. This is not to suggest that extrinsic factors and game-elements do not have any inherent value for volunteern motivation, as the EyeWire team suggest that game elements were highly effective for their 'power players' and had a positive effect in terms of player retention and project efficiency. However, these elements enhance, support or otherwise compliment players' intrinsic interests and the way in which players perceive tasks, rather than replacing such factors. In this way, game-based projects are similar to non-gamified citizen science intiatives, which rely heavily on players' own interests [Wiggins and Crowston 2011]. Nevertheless, there is some evidence in the wider gamification literature to suggest that framing tasks as games can make them more interesting and enjoyable, potentially impacting on players intrinsic motivations. Lieberoth [2015] demonstrates that even the simple act of framing a task as a game is as motivating for participants as the introduction of game elements themselves, although these elements were also suggested to be intrinsically motivating.

Nevertheless, there is a complex interplay between the motivations demonstrated by citizen scientists and we suggest that introducing extrinsic factors to tasks which rely on participants' altruistic or intrinsic motivations can have unexpected effects. In particular, we note the example of EyeWire's Scythe Complete feature, which while of great importance to the project and which was framed in a highly altruistic manner, was nevertheless avoided by players who suggested that it was not fun and offered poor rewards. The team similarly suggested that activities offering higher rewards and point scores receive more attention from players than those which do not, even though in previous studies of EyeWire players have described their motivations as predominantly intrinsic or altruistic [Tinati et al. 2016].

Our findings also contribute to the discussion surrounding the use of competitive and collaborative elements in gamified citizen science projects. While competitive elements appear to be somewhat more common - and in some games, such as EyeWire, form the backbone of gameplay - players as a whole are more engaged by opportunities for cooperation and for social interaction. Nevertheless, different groups of players respond differently to competition and highly active and younger players have been suggested to be more engaged by competitive elements than by collaboration. This suggests a potential trade-off between the needs of a high performing minority and a larger, but less active minority. Nevertheless, mid-performing players may also be motivated by competing with themselves through leaderboard and point features, even if not otherwise engaged through competition with other players and these features may also contribute to 'normalising behaviours', where players feel a (self-driven) social pressure to do their fair share of the work [Preist et al. 2014]. It is unclear from the literature to what extent this phenomenon is unique to gamified and game-based projects. While on the one hand, building relationships with fellow players and scientists and collaborating with the community towards common goals have been identified by players as key factors motivating their participation in citizen science [Nov et al. 2011; Rotman et al. 2014], participants in Galaxy Zoo and other Zooniverse projects did not describe social interaction as a strong source of motivation [Cox et al. 2015; Raddick et al. 2013]. This is perhaps surprising, given that sociality was more common during our project survey among those projects which did not describe themselves as games than those which did.

It should be noted also that gamification is not a 'one size fits all' solution to the problems associated with motivating and engaging citizen scientists in project activities. As demonstrated by our literature review findings, participants who are highly intrinsically motivated and those who are less active or less motivated respond poorly to gamified elements, either because they are perceived as trivialising serious and otherwise motivating tasks, or because they introduce stress and high levels of competitiveness that participants cannot hope to match. However, our findings also suggest that such features also have no effect on average performing users - a group characterised by interview participants four and five as 'lurkers' who respond neither positively nor negatively to the introduction of game elements.

6.2 Use of Gamification in VCS

The most popular game elements used within the surveyed projects were fairly straightforward. Points and leaderboards were commonly used, adding game mechanics to certain tasks and introducing some form of competition. These features were commonly assessed in other studies summarised earlier in this paper. However their complexity in terms of gameplay and engagement is limited, which suggests that there are still opportunities for development.

Our findings show that recommendations and best practices from the literature are not always followed when VCS games are designed and implemented. For example, in spite of the importance associated with altruistic and intrinsic motivational factors in prior studie, these motivations are not necessarily present when considering the features implemented within projects. While the surveyed games demonstrated a high proportion of altruistic framing, such features were absent in non-game projects with gamified elements. VCS games also demonstrate a tendency to make use of competitive elements, in contrast with recommendations from the literature, which suggests that volunteers respond more positively to collaborative and team-based activities.

One possible explanation for the difference between the theory of literature findings and the practise of project design is the relative simplicity of implementing features. The interview findings suggest that some of the features that are most appealing to players, such as narrative framing and progression opportunities, were time-consuming and difficult to implement, due to an incompatibility with VCS contexts — in particular, the lack of a feasible ultimate goal to reach or issue to overcome. Similarly the tendency for reward elements such as points and badges to eventually reduce in value as players exhausted the games reward scheduling, requiring further effort from developers who must implement further mechanisms.

At the same time, more advanced game elements would require both expertise within the project team and an engaged community to demand them. For example, we noted from the existing literature that story-based games could be more interesting for some groups of users [Prestopnik and Tang 2015], yet the use of narrative was rare in the projects that we surveyed (Figure 5). Shaping a game around a narrative could be seen as a considerably bigger development task, predominantly as unlike other 'features' narratives are far more complex and of much more significance - framing both the task and all other features. This was particularly problematic especially when there are restrictions on many of these projects to ensure that they are still scientifically useful—as discussed by the EyeWire team. Due to the relatively low use of narrative so far, it seems that there remains an untapped potential to frame game tasks around some kind of story, which could go some way to reaching those citizens who are not motivated by points or competition. In particular, our findings suggest that these more intrinsically motivated players are currently less well served by VCS games than those who enjoy extrinsic rewards.

EyeWire makes use of different player roles as a way of demonstrating experienced players' advanced status in the game and appealing to various types of motivations. However, in the projects we surveyed, both status and roles were not used to a great extent, and different player roles were not the focus of any of the papers presented in Table 7. Even in EyeWire, the decision to introduce these roles was predominantly motivated by the need to increase the efficency of the project and to pass more time consuming activities into the hands of players. Nevertheless, having these sorts of elements can give players something to work towards, and a reward for doing so, in a similar manner to badges. It is therefore surprising that this element has not been used to a greater extent among other projects, given that it represents a form of progression that is easier to introduce than, for example, levelling systems or narrative arcs. Future research could look into tracking the effect of its introduction. Indeed there are numerous opportunities for carrying out this form of analysis on new elements, where for example A/B testing could be used to demonstrate whether a particular change has a desirable effect. Again, referring back to the papers in Table 7, very little existing literature has taken this approach.

Interestingly, the findings from both the existing literature and the interviews carried out in this study point toward much of the development in VCS games going towards maintaining an existing community of players. The EyeWire team spoke about adding new features for their "power players" and based on feedback from those already playing to sustain the community. This echoes and may partially explain a number of literature findings such as Iacovides et al. [2013] which (by studying EyeWire along with Foldit) found that gamification was useful for sustaining engagement but had less of an effect for attracting new volunteers. As noted, newcomers to VCS projects are predominantly motivated by their own altruistic and intrinsic interests in science, but out findings suggest that many gamified projects lose sight of this. While altruistic framing devices were relatively common, there was a key distinction in non-gaming contexts, where the use of gaming elements occured mutually exclusively with altruistic framing. This may explain the tendency that gamified projects like Old Weather observed among new players leaving the project relatively early [Eveleigh et al. 2013].

Based on these findings, we make the recommendation that VCS games and particularly gamified projects should employ a range of motivational affordances. In their current state, many of the surveyed projects and literature findings suggest that some groups of players are not well represented in terms of the design decisions made by designers to satisfy players — particularly those players who are more intrinsically motivated by the urge to engage with science. Such an effort would have beneficial effects not only for attracting and maintaining player engagement, but also for ensuring the accuracy of project outputs [Mekler et al. 2013].

6.3 Analysis of Gamification Effects

From the literature, it was apparent that although gamification is becoming increasingly discussed within VCS, there is still little coverage of the tracking of specific game element performance. The general reception in the studies presented in Table 7 is that game elements tend to have a positive impact, but in some cases this is challenged, with some suggesting that the use of competitive features—as well as others—can put off certain players [Eveleigh et al. 2013]. However it is unclear whether in general this effect is true, or whether the projects themselves have deeper insights into what impact the different elements do have. From the interviews with the EyeWire team, we found that this analysis was carried out only to a small extent by looking back over player engagement data and whether changes in activity appeared to follow a particular change to the game, rather than tracking the effect of the new element at the time. Instead, change tracking is carried out qualitatively, using insights from the chat feature which allows a two-way feedback mechanism, and can ensure any potentially bad effects are negated quickly. It should be noted nonetheless that comparative analyses of the impact of features - e.g., collaborative vs cooperative activities - were not carried out by the team, in part due to a lack of available resources.

The importance of tracking changes was underlined in the EyeWire interviews when the experience of listening to what the community says, rather than observing what it does, was shown to sometimes be unreliable—especially if a change fundamentally reduces the feeling that the project is a game, or if it causes existing players to start appearing worse. This is a compelling insight, as it suggests that the community around EyeWire *are* there because of the gamified nature of the project, and their motivation is affected when changes subsequently reduce this experience, which would support the findings of numerous papers discussed earlier where points were argued to be engaging features that helped to sustain engagement [Iacovides et al. 2013; Mekler et al. 2013; Siu et al. 2014]. At the same time, this raises questions about the findings of a number of the surveyed publications, which rely on surveys and interviews with VCS players. We suggest that further analysis of these contexts - i.e., motivation and engagement - is required to clarify the effect that more intrinsic and altruistic features have on player engagement, when compared with extrinsic point and reward systems.

Our findings also suggest a number of potential issues in applying game framing and activities to scientific research contexts. While the accuracy of data submissions has been shown to generally be comparable with non-gaming contexts, there remains a question of the development time and effort required to attract and sustain player populations. In EyeWire, the team made frequent note of the various requests made by players and the negative reaction many players had to changing features — regardless of the intention behind the change or the scientific needs which motivated such changes. Similarly, it would suggest that removing game elements that have been used by an established community such as in EyeWire would have a detrimental effect, comparable to those impacts reported in enterprise social networks [Thom et al. 2012]. This trade-off in balancing the needs of game players with the needs of project scientists is an interesting area for further research that was not significantly explored in the literature surveyed.

6.4 Game and Non-game Distinction

Overall, our findings suggest that the distinction between those projects which describe themselves as games and those which do not is not well established. On the one hand, there were certain common trends among the two groups - with non-gaming projects being more likely to include opportunities for collaboration, learning and altruistic framing, while gaming projects were more focused on extrinsic motivations. However, no features were specific to only one type of project and despite not describing themselves as games, many of the 'non-game' projects analysed within our survey featured gamified elements such as points. Moreover, within this non-game category, two clear subtypes emerged - gamified projects, which share many common features with games, with almost exclusively extrinsic features such as points and leaderboards and non-gamified projects, which rely on participants' altruistic intentions.

Yet even among those projects which explicitly describe themselves as games, the extent to which the activities assigned to players resemble a conventional game varies. As the EyeWire team stated, some VCS games better resemble work activities, while others, like EteRNA, very much have the look and feel of a game. The question of how to define a 'game' is therefore a key question within this area — or to paraphrase the interview participants: is it simply enough to describe an activity as a game? Moreover, is it enough to simply add game elements to an inherently work-like task?

Across our three studies, we have found that VCS games are predominantly defined by restraints such as the limited resources available for development and commonalities such as the motivations of volunteers. Nevertheless, in its current state, the genre is not sufficiently defined to warrant specific expectations of what a VCS game might encompass, either in terms of features, or in terms of the activity and experience provided to players. Our interview findings suggest that for EyeWire at least, this was problematic, both for players and for designers, who struggled to meet the expectations of the player base. Similarly, while the surveyed literature included some consideration of the effects of game framing, there was little in the way of exploration of the impact of specific game activities - e.g., puzzle solving - when compared with non-game tasks combined with features such as points.

For this reason, we believe that a key future research area should be ascertaining and exploring the expectations of project volunteers with regard to game and non-game experiences. Given the tendency of EyeWire volunteers to request and even expect new features which were incompatible with citizen science contexts, we identify a need for greater understanding of the inherent opportunities that citizen science provides for game play. Such research findings would be of significant value to the EyeWire team and, we believe, to other researchers looking to make use of gamification to modify crowdsourcing and human computation tasks.

6.5 Limitations and Future Work

Overall, it seems that the attitude towards gamification in citizen science is that it does generally have a positive effect, and as with its use in crowdsourcing, future studies now need to address which specific design choices bring about its success [Morschheuser et al. 2017]. As mentioned above, there is an opportunity to further study the introduction of particular game elements to existing projects, following A/B testing or other means, to determine their overall effect on players. Currently the work done to track the effectiveness of particular elements appears limited, and without deeper insights here there will continue to be a lack of understanding about how best to employ them. Forde et al. [2016] have planned a more in-depth study on the effects of individual game elements, and gamification in general, on the Fylo (Fake Phylo) project, following their original research which produced quantitative data that showed no effect, while qualitative insights suggested the game elements were important. Our findings support their plan, along with their argument for the need to study the area in more depth.

While this study has attempted to establish the current perspectives in existing literature and projects, we appreciate that focusing on one VCS game for the interviews only provides a limited window onto the wider VCS scene. However we believe that the views and experiences of the EyeWire team reflect many of the discussions put forward in existing literature. Furthermore, while we have analysed existing research for insights into the effects of gamification, we recognise that a large-scale study across multiple projects and looking at longitudinal engagement data has yet to be conducted.

In addition, one remaining area of ambiguity is in the impact of game elements on players intrinsic motivations and the effect this has on the recruitment of participants. Our findings suggest that game elements have little impact on a players' initial decision to contribute to a project and that this decision is made based on a players' own intrinsic interests. However, there are some suggestions in the literature that extrinsic factors and game elements have a moderate effect on players' intrinsic motivations. It is also notable that due to the nature of VCS, there has so far been little opportunity for A-B testing to compare the importance of science and scientific framing with the impact of game elements and game-like framing. We nevertheless suggest that this is an interesting area for further study which could be incorporated into the research process when comparing the effectiveness of individual game elements.

7 FINAL REMARKS AND CONCLUSIONS

This paper has provided a range of insights into the use of gamification within virtual citizen science projects. While success in citizen science has been discussed before, there has yet to be the same examination of what constitutes effective use of game elements in these projects. We provide an overview of the literature reporting on the use of individual elements and projects, and summarise the current landscape with regards to gamified VCS. We explored these insights further through the interviews with the EyeWire project team, and found that there appears to be significant focus on maintaining and satisfying existing communities, rather than reaching new players, and that development of such features that would motivate new users is hindered by resource limitations compared to traditional game studios. Our findings have demonstrated that the impact of gamification and game elements on VCS projects is complex - while active and engaged players are highly motivated by extrinsic and game factors, those players who are more driven by intrinsic motivations perceive game elements negatively. Moreover, while extrinsic factors may support long-term participation, intrinsic motivations and altruism remain important drivers for initial participation and the switch to long-term engagement patterns within VCS. For this reason, game elements should not be used alone, but should rather be introduced alongside framing devices that support other motivations such as opportunities for learning and indications of the value of players' contributions for scientific research. We envisage that the use of game elements in VCS will continue to rise, however it is clear that more robust tracking and analysis mechanisms are required in order to truly understand the impact of each particular feature.

REFERENCES

David P Anderson, Jeff Cobb, Eric Korpela, Matt Lebofsky, and Dan Werthimer. 2002. SETI@ home: an experiment in public-resource computing. *Commun. ACM* 45, 11 (2002), 56–61.

Andrés Francisco Aparicio, Francisco Luis Gutiérrez Vela, José Luis González Sánchez, and José Luis Isla Montes. 2012. Analysis and Application of Gamification. In *Proceedings of the 13th International Conference on InteracciÓN Persona-Ordenador (INTERACCION '12)*. ACM, New York, NY, USA, Article 17, 2 pages. https://doi.org/10.1145/2379636.2379653

Avinoam Baruch, Andrew May, and Dapeng Yu. 2016. The motivations, enablers and barriers for voluntary participation in an online crowdsourcing platform. *Computers in Human Behavior* 64 (2016), 923–931.

Dr Blohm, Ivo and Prof Dr Leimeister, Jan M. 2013. Gamification. *Business Information Systems Engineering* 5, 4 (08 2013), 275–278. https://search.proquest.com/docview/1412812105?accountid=13963 Copyright - Springer Fachmedien Wiesbaden 2013; Document feature - ; Tables; Last updated - 2014-01-22; SubjectsTermNotLitGenreText - United States–US.

John Bohannon. 2014. Online Video Game Plugs Players Into Remote-Controlled Biochemistry Lab. *Science* 343, 6170 (2014), 475–475. https://doi.org/10.1126/science.343.6170.475 arXiv:http://science.sciencemag.org/content/343/6170/475.full.pdf

Anne Bowser, Derek Hansen, Yurong He, Carol Boston, Matthew Reid, Logan Gunnell, and Jennifer Preece. 2013. Using gamification to inspire new citizen science volunteers. In *Proceedings of the first international conference on gameful design, research, and applications*. ACM, 18–25.

A. Carlier, A. Salvador, F. Cabezas, X. Giro-i Nieto, V. Charvillat, and O. Marques. 2016. Assessment of crowdsourcing and gamification loss in user-assisted object segmentation. *Multimedia Tools and Applications* 75, 23 (2016), 15901–15928. https://doi.org/10.1007/s11042-015-2897-6

Chen Chen, Paweł W. Woźniak, Andrzej Romanowski, Mohammad Obaid, Tomasz Jaworski, Jacek Kucharski, Krzysztof Grudzień, Shengdong Zhao, and Morten Fjeld. 2016. Using Crowdsourcing for Scientific Analysis of Industrial Tomographic Images. *ACM Trans. Intell. Syst. Technol.* 7, 4, Article 52 (July 2016), 25 pages. https://doi.org/10.1145/2897370

C. Chung, A. Kadan, Y. Yang, A. Matsuoka, J. Rubin, and M. Chechik. 2017. The impact of visual load on performance in a human-computation game. *ACM International Conference Proceeding Series* Part F130151 (2017). https://doi.org/10.1145/3102071.3106358

Joe Cox, Eun Young Oh, Brooke Simmons, Chris Lintott, Karen Masters, Gary Graham, Anita Greenhill, and Kate Holmes. 2015. How is success defined and measured in online citizen science? A case study of Zooniverse projects. (2015).

Kevin Crowston and Isabelle Fagnot. 2008. The motivational arc of massive virtual collaboration. In *Proceedings of the IFIP WG*, Vol. 9.

Kevin Crowston and Nathan R Prestopnik. 2013. Motivation and data quality in a citizen science game: A design science evaluation. In *System Sciences (HICSS), 2013 46th Hawaii International Conference on*. IEEE, 450–459.

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

Edward L Deci, Richard Koestner, and Richard M Ryan. 1999. A Meta-Analytic Review of Experiments Examining the Effects of Extrinsic Rewards on Intrinsic Motivation. *Psychological Bulletin* 125, 6 (1999), 627–6.

Sebastian Deterding. 2011. Situated motivational affordances of game elements: A conceptual model. In *Gamification: Using game design elements in non-gaming contexts, a workshop at CHI.*

Sebastian Deterding. 2012. Gamification: designing for motivation. *interactions* 19, 4 (2012), 14–17.

Sebastian Deterding, Dan Dixon, Rilla Khaled, and Lennart Nacke. 2011a. From game design elements to gamefulness: defining gamification. In *Proceedings of the 15th international academic MindTrek conference: Envisioning future media environments.* ACM, 9–15.

Sebastian Deterding, Miguel Sicart, Lennart Nacke, Kenton O'Hara, and Dan Dixon. 2011b. Gamification. using game-design elements in non-gaming contexts. In *CHI'11 Extended Abstracts on Human Factors in Computing Systems.* ACM, 2425–2428.

Nicolas Ducheneaut, Robert J Moore, and Eric Nickell. 2007. Virtual "third places": A case study of sociability in massively multiplayer games. *Computer Supported Cooperative Work (CSCW)* 16, 1-2 (2007), 129–166.

David Easley and Arpita Ghosh. 2016. Incentives, Gamification, and Game Theory: An Economic Approach to Badge Design. *ACM Trans. Econ. Comput.* 4, 3, Article 16 (June 2016), 26 pages. https://doi.org/10.1145/2910575

Alexandra Eveleigh, Charlene Jennett, Ann Blandford, Philip Brohan, and Anna L Cox. 2014. Designing for dabblers and deterring drop-outs in citizen science. In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems.* ACM, 2985–2994.

Alexandra Eveleigh, Charlene Jennett, Stuart Lynn, and Anna L Cox. 2013. "I want to be a Captain! I want to be a Captain!": Gamification in the Old Weather Citizen Science Project. In *Proceedings of the first international conference on gameful design, research, and applications.* ACM, 79–82.

Seamus F. Forde, Elisa D. Mekler, and Klaus Opwis. 2016. Informational, but Not Intrinsically Motivating Gamification': Preliminary Findings. In *Proceedings of the 2016 Annual Symposium on Computer-Human Interaction in Play Companion Extended Abstracts (CHI PLAY Companion '16).* ACM, New York, NY, USA, 157–163. https://doi.org/10.1145/2968120.2987738

D.H.-L. Goh, R.P. Ang, A.Y.K. Chua, and C.S. Lee. 2010. Evaluating game genres for tagging images, In Proceedings of the 6th Nordic Conference on Human-Computer Interaction: Extending Boundaries. *NordiCHI 2010: Extending Boundaries - Proceedings of the 6th Nordic Conference on Human-Computer Interaction* (2010), 659–662. https://doi.org/10.1145/1868914.1868998

GG Graham, J Cox, Brooke Simmons, Chris Lintott, Karen Masters, A Greenhill, and K Holmes. 2015. How is success defined and measured in online citizen science: a case study of Zooniverse projects. *Computing in science and engineering* 99 (2015), 22.

A. Greenhill, K. Holmes, J. Woodcock, C. Lintott, B.D. Simmons, G. Graham, J. Cox, E.Y. Oh, and K. Masters. 2016. Playing with science: Exploring how game activity motivates users participation on an online citizen science platform. *Aslib Journal of Information Management* 68, 3 (2016), 306–325. https://doi.org/10.1108/AJIM-11-2015-0182

Juho Hamari. 2013. Transforming homo economicus into homo ludens: A field experiment on gamification in a utilitarian peer-to-peer trading service. *Electronic commerce research and applications* 12, 4 (2013), 236–245.

Juho Hamari, Jonna Koivisto, and Harri Sarsa. 2014. Does gamification work?–a literature review of empirical studies on gamification. In *2014 47th Hawaii International Conference on System Sciences.* IEEE, 3025–3034.

Jeff Howe. 2006. The rise of crowdsourcing. *Wired magazine* 14, 6 (2006), 1–4.

Shih-Wen Huang and Wai-Tat Fu. 2012. Systematic analysis of output agreement games: Effects of gaming environment, social interaction, and feedback. *Urbana* 51 (2012), 61801.

Kai Huotari and Juho Hamari. 2012. Defining gamification: a service marketing perspective. In *Proceeding of the 16th International Academic MindTrek Conference.* ACM, 17–22.

Ioanna Iacovides and Anna L Cox. 2015. Moving beyond fun: Evaluating serious experience in digital games. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems.* ACM, 2245–2254.

Ioanna Iacovides, Charlene Jennett, Cassandra Cornish-Trestrail, and Anna L Cox. 2013. Do games attract or sustain engagement in citizen science?: a study of volunteer motivations. In *CHI'13 Extended Abstracts on Human Factors in Computing Systems.* ACM, 1101–1106.

Charlene Jennett, Laure Kloetzer, Daniel Schneider, Ioanna Iacovides, Anna Cox, Margaret Gold, Brian Fuchs, Alexandra Eveleigh, Kathleen Methieu, Zoya Ajani, et al. 2016. Motivations, learning and creativity in online citizen science. *Journal of Science Communication* 15, 3 (2016).

Ippokratis Kapenekakis and Konstantinos Chorianopoulos. 2017. Citizen science for pedestrian cartography: collection and moderation of walkable routes in cities through mobile gamification. *Human-centric Computing and Information Sciences* 7, 1 (2017), 10.

Alexander Kawrykow, Gary Roumanis, Alfred Kam, Daniel Kwak, Clarence Leung, Chu Wu, Eleyine Zarour, Luis Sarmenta, Mathieu Blanchette, Jérôme Waldispühl, et al. 2012. Phylo: a citizen science approach for improving multiple sequence alignment. *PloS one* 7, 3 (2012), e31362.

Jeffrey Laut, Francesco Cappa, Oded Nov, and Maurizio Porfiri. 2016. Increasing citizen science contribution using a virtual peer. *Journal of the Association for Information Science and Technology* (2016), n/a–n/a. https://doi.org/10.1002/asi.23685

Andreas Lieberoth. 2015. Shallow gamification testing psychological effects of framing an activity as a game. *Games and Culture* 10, 3 (2015), 229–248.

Chris J Lintott, Kevin Schawinski, Anže Slosar, Kate Land, Steven Bamford, Daniel Thomas, M Jordan Raddick, Robert C Nichol, Alex Szalay, Dan Andreescu, et al. 2008. Galaxy Zoo: morphologies derived from visual inspection of galaxies from the Sloan Digital Sky Survey. *Monthly Notices of the Royal Astronomical Society* 389, 3 (2008), 1179–1189.

Jane McGonigal. 2011. *Reality is broken: Why games make us better and how they can change the world.* Penguin.

Elisa D Mekler, Florian Brühlmann, Klaus Opwis, and Alexandre N Tuch. 2013. Disassembling gamification: the effects of points and meaning on user motivation and performance. In *CHI'13 extended abstracts on human factors in computing systems.* ACM, 1137–1142.

Benedikt Morschheuser, Juho Hamari, and Jonna Koivisto. 2016. Gamification in crowdsourcing: A review. In *System Sciences (HICSS), 2016 49th Hawaii International Conference on.* IEEE, 4375–4384.

Benedikt Morschheuser, Juho Hamari, Jonna Koivisto, and Alexander Maedche. 2017. Gamified Crowdsourcing: Conceptualization, Literature Review, and Future Agenda. *International Journal of Human-Computer Studies* (2017).

Oded Nov, Ofer Arazy, and David Anderson. 2011. Dusting for science: motivation and participation of digital citizen science volunteers. In *Proceedings of the 2011 iConference*. ACM, 68–74.

M.K. Pedersen, N.R. Rasmussen, J. Sherson, and R.V. Basaiawmoit. 2017. Leaderboard effects on player performance in a citizen science game. *Proceedings of the 11th European Conference on Games Based Learning, ECGBL 2017* (2017), 531–537.

M. Ponti, T. Hillman, and I. Stankovic. 2015. Science and gamification: The odd couple?, In Proceedings of the 2015 Annual Symposium on Computer-Human Interaction in Play. *CHI PLAY 2015 - Proceedings of the 2015 Annual Symposium on Computer-Human Interaction in Play* (2015), 679–684. https://doi.org/10.1145/2793107.2810293

Catherine Pope, Sue Ziebland, Nicholas Mays, et al. 2000. Analysing qualitative data. *Bmj* 320, 7227 (2000), 114–116.

C. Preist, E. Massung, and D. Coyle. 2014. Competing or aiming to be average? Normification as a means of engaging digital volunteers. *Proceedings of the ACM Conference on Computer Supported Cooperative Work, CSCW* (2014), 1222–1233. https://doi.org/10.1145/2531602.2531615

Nathan Prestopnik, Kevin Crowston, and Jun Wang. 2014. Exploring data quality in games with a purpose. *iConference 2014 Proceedings* (2014).

N. Prestopnik, K. Crowston, and J. Wang. 2017. Gamers, citizen scientists, and data: Exploring participant contributions in two games with a purpose. *Computers in Human Behavior* 68 (mar 2017), 254–268. https://doi.org/10.1016/j.chb.2016.11.035

Nathan R Prestopnik and Jian Tang. 2015. Points, stories, worlds, and diegesis: comparing player experiences in two citizen science games. *Computers in Human Behavior* 52 (2015), 492–506.

M Jordan Raddick, Georgia Bracey, Pamela L Gay, Chris J Lintott, Carie Cardamone, Phil Murray, Kevin Schawinski, Alexander S Szalay, and Jan Vandenberg. 2013. Galaxy Zoo: Motivations of Citizen Scientists. *Astronomy Education Review* 12, 1 (2013).

Neal Reeves, Ramine Tinati, Sergej Zerr, Max G Van Kleek, and Elena Simperl. 2017. 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*. ACM, 2137–2152.

Francesco Restuccia, Sajal K. Das, and Jamie Payton. 2016. Incentive Mechanisms for Participatory Sensing: Survey and Research Challenges. *ACM Trans. Sen. Netw.* 12, 2, Article 13 (April 2016), 40 pages. https://doi.org/10.1145/2888398

Markus Rokicki, Sergej Zerr, and Stefan Siersdorfer. 2016. Just in Time: Controlling Temporal Performance in Crowdsourcing Competitions. In *Proceedings of the 25th International Conference on World Wide Web (WWW '16)*. International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, Switzerland, 817–827. https://doi.org/10.1145/2872427.2883075

Dana Rotman, Jen Hammock, Jenny Preece, Derek Hansen, Carol Boston, Anne Bowser, and Yurong He. 2014. Motivations affecting initial and long-term participation in citizen science projects in three countries. *iConference 2014 Proceedings* (2014).

Dana Rotman, Jenny Preece, Jen Hammock, Kezee Procita, Derek Hansen, Cynthia Parr, Darcy Lewis, and David Jacobs. 2012. Dynamic Changes in Motivation in Collaborative Citizen-science Projects. In *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work (CSCW '12)*. ACM, New York, NY, USA, 217–226. https://doi.org/10.1145/2145204.2145238

David Rudrum. 2005. From narrative representation to narrative use: Towards the limits of definition. *Narrative* 13, 2 (2005), 195–204.

M. Sabou, K. Bontcheva, A. Scharl, and M. FÃűls. 2013. Games with a purpose or mechanised labour? A comparative study, In Proceedings of the 13th International Conference on Knowledge Management and Knowledge Technologies. *ACM International Conference Proceeding Series*, Article 19 (2013), Article 19, 8 pages. https://doi.org/10.1145/2494188.2494210

Katie Seaborn and Deborah I Fels. 2015. Gamification in theory and action: A survey. *International Journal of Human-Computer Studies* 74 (2015), 14–31.

K. Siu, M. Guzdial, and M.O. Riedl. 2017. Evaluating singleplayer and multiplayer in human computation games. *ACM International Conference Proceeding Series* Part F130151 (2017). https://doi.org/10.1145/3102071.3102077

K. Siu and M.O. Riedl. 2016. Reward systems in human computation games, In CHI PLAY 2016: PROCEEDINGS OF THE 2016 ANNUAL SYMPOSIUM ON COMPUTER-HUMAN INTERACTION IN PLAY. *CHI PLAY 2016 - Proceedings of the 2016 Annual Symposium on Computer-Human Interaction in Play* (2016), 266–275. https://doi.org/10.1145/2967934.2968083

Kristin Siu, Alexander Zook, and Mark O Riedl. 2014. Collaboration versus competition: design and evaluation of mechanics for games with a purpose. In *Proceedings of the 9th International Conference on the Foundations of Digital Games*.

S. Thaler, E. Simperl, and S. WÃűlger. 2012. An experiment in comparing human-computation techniques. *IEEE Internet Computing* 16, 5 (sep-oct 2012), 52–58. https://doi.org/10.1109/MIC.2012.67

Sarah-Kristin Thiel. 2016. A Review of introducing Game Elements to e-Participation. In *E-Democracy and Open Government (CeDEM), Conference for*. IEEE, 3–9.

Sarah-Kristin Thiel, Michaela Reisinger, Kathrin Röderer, and Peter Fröhlich. 2016. Playing (with) Democracy: A Review of Gamified Participation Approaches. *JeDEM-eJournal of eDemocracy and Open Government* 8, 3 (2016), 32–60.

Jennifer Thom, David Millen, and Joan DiMicco. 2012. Removing gamification from an enterprise SNS. In *Proceedings of the acm 2012 conference on computer supported cooperative work*. ACM, 1067–1070.

Ramine Tinati, Markus Luczak-Roesch, Elena Simperl, and Wendy Hall. 2016. Because science is awesome: studying participation in a citizen science game. In *Proceedings of the 8th ACM Conference on Web Science*. ACM, 45–54.

R. Tinati, M. Luczak-Roesch, E. Simperl, and W. Hall. 2017a. An investigation of player motivations in Eyewire, a gamified citizen science project. *Computers in Human Behavior* 73 (aug 2017), 527–540. https://doi.org/10.1016/j.chb.2016.12.074

Ramine Tinati, Elena Simperl, and Markus Luczak-Roesch. 2017b. To help or hinder: Real-time chat in citizen science. In *11th International Conference on Web and Social Media*. https://eprints.soton.ac.uk/406181/

Ramine Tinati, Max Van Kleek, Elena Simperl, Markus Luczak-Rösch, Robert Simpson, and Nigel Shadbolt. 2015. Designing for Citizen Data Analysis: A Cross-Sectional Case Study of a Multi-Domain Citizen Science Platform. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. ACM, New York, NY, USA, 4069–4078. https://doi.org/10.1145/2702123.2702420

A. Wiggins and K. Crowston. 2011. From Conservation to Crowdsourcing: A Typology of Citizen Science. In *2011 44th Hawaii International Conference on System Sciences*. 1–10. https://doi.org/10.1109/HICSS.2011.207

A SUMMARY OF LITERATURE REVIEW PUBLICATIONS

Table 7. An overview of the findings in the literature about gamification use in citizen science projects.

Study	Project	Method	Game element or features assessed	Psychological Outcome	Behavioural Outcome	Positive / negative impact of gamification
Bowser et al. [2013]	Biotracker	Field Study (N=71)	Badges; leaderboards	Game elements increase motivation to participate.		Positive
Carlier et al. [2016]	Ask'n'Seek	Web-based Study (N=162)	Game Framing		Competition leads to reduced level of contribution; Higher level of accuracy (relative to paid crowdsourcing)	Mixed
Chung et al. [2017]	Matchmakers	Web-based Study (37)	Visual Load		Cumulative Score and successful game completions higher with greater visual load	Positive
Crowston and Prestopnik [2013]	Happy Match / Citizen Sort	Web-based Experiment (N=200)	Points;	Some reported motivation to play the game.		Neutral
Curtis [2015]	Foldit	Survey (N=37); Interview (N=10)	Points; Competition; Cooperation; Sociability; Tutorial;	Participants believed they were making an important contribution to science. Little motivation from points and ranking mechanics.		Mixed: positive for less-traditional elements such as chat, less so for points and leaderboards.
Eveleigh et al. [2013]	Old Weather	Survey (N=545); Interview (N=18)	Progression/Status; Leaderboards; Competition	Motivates some volunteers.	Competition leads to casual participants leaving.	Mixed
Forde et al. [2016]	Fylo ('Fake Phylo')	Web-based Experiment (N=67)	Points, score bar, accuracy	No significant differences from gamified version in terms of intrinsic motivation, competence and autonomy.	Game elements were informational and useful for adapting behaviour.	Mixed
Goh et al. [2010]	Bespoke Experiment	Lab Study (N=66)	Genre - Cooperative vs Competitive		Competitive activities lead to significantly higher numbers of tags and tag quality	Mixed
Greenhill et al. [2016]	Zooniverse	Participant Observation	Extrinsic Rewards	Participants motivated by both intrinsically interesting and scientifically valuable tasks and extrinsically rewarded 'just for fun' tasks		Mixed

Huang and Fu [2012]	Bespoke Experiment	Web-based Experiment (N=150)	Game Framing; Feedback; Social Interaction		Game framing and feedback both lead to increase in contribution levels - particularly when combined. Social interaction increases participation levels, but to slightly lesser extent	Mixed
Iacovides et al. [2013]	Foldit, EyeWire	Interview (N=8)	Cooperation; sociability; leaderboards; Points	Participants wanted to feel they were making a difference; felt that evidence of project progress was rewarding; chart and forums increased sense of community.	Points and leaderboards extend sessions; Teamplay led to continued engagement.	Positive for sustaining engagement, no effect for attracting new volunteers
Iacovides and Cox [2015]	Medical Student Errors; Nurse's Dilemma; Patient Panic; St. Error Hospital	Player Evaluation (N=12)	Game framing	serious activities as games motivates players through development of empathy and emotional connection		Positive
Kapenekakis and Chorianopoulos [2017]	Bespoke mobile app	Field Study (N=13)	Points	Some users motivated to find new routes, and moderate bad behaviour.	Many players didn't add road tags, despite generous offer of points.	Generally positive
Laut et al. [2016]	Brooklyn Atlantis	Web Study (N=53,090)	Competition (through virtual peer)		Participants adjust performance based on that of their virtual peer, and contributions can be increased.	Positive
Lieberoth [2015]	Bespoke Experiment	Lab Study (N=90)	Framing	Framing tasks as games increases participant motivation even regardless of game features		Positive
Mekler et al. [2013]	Bespoke experiment	Web-based Experiment (N=172)	Points; framing	Points and framing both increased intrinsic motivation.	Using points generated more tags; more time spent per tag when meaningful framing was combined with points; framing increased tag quality.	Positive
Pedersen et al. [2017]	Quantum Moves	Web-based Experiment (N=4553)	Leaderboards		No statistically significant impact on player retention or accuracy of submissions	Mixed
Ponti et al. [2015]	FoldIt, Galaxy Zoo	Interviews (N=24); Content Analysis (N=650,000)	Points; Status	Point scores motivate participation for some players - others demotivated due to use of scripts to automate processes	Public status announcements lead to cheating behaviour and low quality submissions	Mixed

Preist et al. [2014]	Close The Door	Field Study	Points; Leaderboards	Leaderboards most motivating for top players. Average players motivated by normification and personal pressure to contribute to work		Mixed
Prestopnik et al. [2014]	Forgotten Island, Happy Match	Web-based Experiment (N=900)	Points; narrative		Story-based game introduced more cheating behaviour, without severely impacting volunteer accuracy. Volunteer retention better in points-based game.	Generally positive
Prestopnik and Tang [2015]	Forgotten Island, Happy Match	Lab Experiment (N=27)	Points; narrative	Story-based game more interesting; points-based games that focus on the science less appealing to those intrinsically motivated.		Positive
Prestopnik et al. [2017]	Forgotten Island, Happy Match	Web-based Study (N=4174)	Narrative, Game Framing		Narrative-based game leads to reduced participant retention; No significant difference in data quality, in spite of cheating in narrative-based game	Mixed
Sabou et al. [2013]	Climate Quiz	Web-based Study (N=648)	Game Framing		Game interface leads to significant increase in task skipping and significant increase in project completion time; Game framing increases accuracy of submissions compared to paid microtask crowdsourcing	Mixed
Siu et al. [2014]	Cabbage Quest (GWAP)	Web-based Experiment (N=118)	Points; competition; cooperation	Participants reported greatest engagement with competitive option.	Collaboration focus increases accuracy, whereas competition increases speed.	Positive
Siu and Riedl [2016]	Cafe Flour Sack	Web-based Experiment (N=78)	Extrinsic Rewards; Points; Leaderboards	Players more motivated when allowed to choose rewards	When choosing rewards, player accuracy, number of contributions and contribution speeds greatly increase	Positive
Siu et al. [2017]	Gwario	Lab-based Study (N=64), Survey (N=64)	Competitive vs Cooperative gaming	Players more motivated by competitive gaming	Cooperative games lead to slight increase in submission accuracy but significant increase in task completion time	Mixed
Thaler et al. [2012]	OntoPronto	Web-based Study (N=270)	Game Framing		Game-based activity leads to increase in number of contributions but reduced submission accuracy	Mixed

Tinati et al. [2017b]	EyeWire	Survey (N=1365)	Sociability real-time chat)		Players establish collaborative mechanism without formal features for these activities.	Positive
Tinati et al. [2017a]	EyeWire	Survey (N=1505)	Learning, 3D Interface	Players predominantly motivated by altruistic desire to assist science, intrinsic interests in science and having fun		Positive

B SURVEYED PROJECTS

Table 8. Projects Surveyed for Game Element Usages. * = Project no longer available at time of publication

Project	**URL**	**Domain**	**Game**
Agent Exoplanet	https://lco.global/agentexoplanet/	Astrophysics	No
Age Guess	https://www.ageguess.org/	Biology	Yes
ARTigo	https://www.artigo.org/	Art History	Yes
Asteroid Vesta	https://cosmoquest.org/x/science/vesta/	Astrophysics	No
Asteroid Zoo	https://www.asteroidzoo.org/	Astrophysics	No
Brooklyn Atlantis	https://cb.engineering.nyu.edu/participate	Environmental Monitoring	No
Cities at Night	https://crowdcrafting.org/project/nightcitiesiss/	Astrophysics	No
Disk Detective	https://www.diskdetective.org/	Astrophysics	No
EteRNA	http://www.eternagame.org/	Biology	Yes
Explore the Seafloor	http://ets.wessexarch.co.uk/	Biology	No
EyeWire	https://eyewire.org/	Neurology	Yes
FoldIt	http://fold.it	Biology	Yes
Galaxy Explorer	https://www.galaxyexplorer.net.au/	Astrophysics	No
Galaxy Zoo	https://www.zooniverse.org/projects/zookeeper/galaxy-zoo	Astrophysics	No
Geo-Wiki	https://www.geo-wiki.org/	Environmental Monitoring	Yes
Mark2Cure	https://mark2cure.org/	Biology	No

McMaster Postcard Project	https://postcards.mcmaster.ca/	History	No
Moon Mappers	https://cosmoquest.org/x/science/moon/	Astrophysics	No
Old Weather	https://www.oldweather.org/	Climatology	No
Phylo	http://phylo.cs.mcgill.ca/	Biology	Yes
Play to Cure: Genes in Space	http://www.genes-in-space.org.uk/*	Immunology	Yes
Reading Nature's Library	https://www.zooniverse.org/projects/mzfasdg2/reading-natures-library	Various	No
Reverse the Odds	https://www.oncology.ox.ac.uk/page/reverse-odds	Immunology	Yes
Smithsonian Transcription Centre	https://transcription.si.edu/	Various	No
Smorball	http://smorballgame.org/	Botany, History, Zoology	Yes
Socientize	http://www.socientize.eu/?q=eu	Various	Yes
Stupid Robot	http://www.tiltfactor.org/game/stupid-robot/	Humanities	Yes
The Skynet	http://www.theskynet.org/	Astrophysics	No
VerbCorner	https://archive.gameswithwords.org/VerbCorner/	Linguistics	Yes
Weather Detective	http://www.weatherdetective.net.au/*	Climatology	No
Wildsense Tigers	http://www.wildsense.org/	Ecology/Conservation	Yes

C INTERVIEW QUESTIONS

C.1 Participant Introduction

(1) Could you begin by introducing yourself and your role in the team?
(2) What is your interaction with other members of the team like?

C.2 Use and introduction of game elements

(1) What are the most important game elements that you use in your project?
(2) How effective do you think these elements have been?
(3) When introducing new features, do you base your development lifecycle on features you know work in other games, or your own insights about how players use your own platform?

(a) Please discuss how has this changed over time?
(b) Are there any features or elements that youâĂŹve considered introducing but subsequently havenâĂŹt done so and if so, why?

C.3 Feedback on new features

(1) After implementing a new feature, what is your process for receiving feedback on it from the community?
(a) How do you review and act upon that feedback?
(b) What is the relationship between data analytics on the platform and the way in which you re-engineer features, compared to when you speak to the community?

C.4 Effect Analysis

(1) When looking at effects on user engagement or motivation, do you study this longitudinally, or just at a single point in time?
(2) When things go wrong, or badly, what process do you follow in order to pinpoint exactly what caused the problem?

C.5 Success in Citizen Science and Games

(1) What do you think makes a CS project successful?
(2) What are your measures for success?
(3) What do you think makes a game successful?
(4) Do all the aspects you've described previously apply to creating a successful CS game? If not, why?