Coordinating Advanced Crowd Work: Extending Citizen Science

Kevin Crowston
Syracuse University
crowston@syr.edu

Erica Michelle Mitchell
Syracuse University
emmitc01@syr.edu

Carsten Østerlund
Syracuse University
costerlu@syr.edu

Abstract

This paper presents a case study of an online citizen science project that attempted to involve volunteers in tasks with multiple dependencies including analyzing bulk data as well as interpreting data and writing a paper for publication. Tasks with more dependencies call for more elaborate coordination mechanisms. However, the relationship between the project and its volunteers limits how work can be coordinated. Contrariwise, a mismatch between dependencies and available coordination mechanisms can lead to performance problems, as were seen in the case. The results of the study suggest recommendations for design of online citizen science projects for advanced tasks.

Introduction

The past decade has seen a rapid growth in the number of online citizen science projects. Most of these projects involve the crowd in rather simple tasks (i.e., microtasking). However, studies of citizen science volunteers suggest that many are motivated by the opportunity to contribute to real science (Raddick et al. 2010; Rotman et al. 2014) and by recognition for such contributions

(Rotman et al. 2012). Accordingly, some sponsors of citizen science projects seek to involve volunteers more deeply in the science of the project: Not just collecting or processing data, but also taking part in further data analysis and even paper writing (Newman et al. 2012).

Efforts to involve volunteers in more advanced tasks also are viewed as important to demonstrate that citizen science is not just crowdsourcing without pay, i.e., an exploitation of the volunteers by project scientists. To be fair to the volunteers, project scientists need to give back (Riesch and Potter 2014, Owens 2016), and expanding access to science is one way to do so. Allowing participants to see and talk about the data is a first step in expanding access (Woodcock et al. 2017).

More advanced science tasks often comprise many interdependent pieces. To successfully include volunteers more deeply in scientific research requires careful consideration of the kinds of project management needed, i.e., how to coordinate contributions to tasks with more dependencies. Accordingly, this paper addresses the following research question: *What coordination challenges do citizen science volunteers face when undertaking work with a high level of task dependency?* The answer to this question has implications both for the theory of citizen science volunteer work and for the practice of including volunteers in more advanced science tasks.

To answer the research question and to provide a model for analyzing dependencies and coordination needs, this paper presents a case study of work in the Galaxy Zoo Quench project. Sponsored by the Zooniverse, Quench volunteers were invited to analyze data (from initial coding through hypothesis testing) and then to use the results of that analysis to write an academic paper in collaboration with the project scientists. The topic of the Quench project was "quenched" galaxies, that is, galaxies that have ceased star formation. Galaxies can quench for different reasons, and understanding why different kinds of galaxies quench can shed light on the processes of galaxy

evolution. The plan was to code a collection of quenched galaxies for various properties and then compare those galaxies to a matched sample of unquenched galaxies (a control group) to identify the distinctive properties of the quenched galaxies. Volunteers would classify the galaxies, as in other citizen science projects; conduct data analysis on the sample and compare it to the control group; and co-author a professional journal article.

The original Galaxy Zoo project already had realized great success involving volunteers in classifying galaxies (the first step of the Quench project). The capability of the volunteers to do original work had seemingly been proven by discoveries such as Hanny's Voorwerp, a novel astronomical object identified by a citizen scientist (Lintott et al. 2009). Furthermore, volunteers had been observed engaging in their own analyses of project data (the planned intermediate steps of the Quench project), posting questions and results to the discussion boards (Tinati et al. 2015; Bonney et al. 2009) and even participating individually in further research. It thus seemed feasible and interesting to involve volunteers in a new scientific collaboration ranging from data analysis to publication (which would have been new for the volunteers). However, this process (writing in particular) also exposed new kinds of coordination challenges for the volunteers. Our analysis of these challenges led us to predict that the volunteers would encounter difficulties in writing the proposed paper. Unexpectedly, they encountered difficulties earlier in the process, with the result that the volunteers never started writing the paper, as we will explain through our case analysis.

Coordination Theory

We draw on coordination theory to explore the coordination challenges associated with engaging citizen science volunteers in advanced science tasks. The case that we examine is rich and can be viewed from numerous perspectives, but coordination theory provides insight into the challenges faced by a distributed group of volunteers trying to work together. In this section, we first introduce

the topic of coordination and present the fundamentals of coordination theory. We next present an analysis of the work of citizen science projects such as Galaxy Zoo, because this kind of work was the foundation of the planned project. Then we describe the process of writing an article, to explore the nature of dependencies and coordination that would arise in the later phases of the Quench project.

Coordination, defined as "managing dependencies between activities" (Malone and Crowston 1994), is a central feature of collective action. Coordination theory (Malone and Crowston 1994) synthesizes contributions from different disciplines to develop a systematic approach to the study of coordination. The theory provides a set of concepts that help make sense of data and which make observed events meaningful, rather than a set of laws that make predictions about relationships between concepts. In other words, coordination theory is a pattern model (Kaplan 1964/1998, p. 327) that seeks to explain phenomena by showing how they fit a known pattern. We use the concepts from this theory (highlighted in italics in the discussion below) to unpack the events of the case and to diagnose the problems faced by the volunteers. Citizen science project managers could similarly analyze work that they are considering asking volunteers to perform to identify and head off potential coordination difficulties.

Concepts of coordination theory

Malone and Crowston (1994) analyzed group action in terms of *actors* performing interdependent *tasks* to achieve some *goal*; i.e., in an organizational process (Crowston 1997; Crowston and Osborn 2003). These *tasks* might require or create various *resources*. For example, in the case of writing a scientific paper, *actors* include the authors and various members of the research team. *Tasks* include collecting data, performing analyses, and writing and revising a manuscript. *Resources* include data, analysis reports, and the analysts' and authors' time and effort. In this view,

actors in collective action face *coordination problems* arising from *dependencies* that constrain how *tasks* can be performed. Studying coordination thus means analyzing the *dependencies* that emerge among the *tasks* and identifying how those *dependencies* are managed. Note that coordination theory focuses on the dependences between tasks, but does not analyze the details of the tasks themselves, e.g., their difficulty or diversity.

In contrast to other theories that consider dependencies among actors, coordination theory posits dependencies as occurring: 1) between a task and a resource and 2) among multiple tasks and a resource. The dependencies between a task and a resource (type 1) are shown in Figure 1. These dependencies exist when a task uses or creates a resource. For example, a data analysis task uses data that has been collected and preprocessed and creates analysis reports. An important kind of resource is the time and effort of an actor (human or otherwise) who can work on the task.

Shared use of resources by different tasks (type 2) can in turn lead to dependencies between the tasks that use or create the resource. These dependencies come in three kinds, as shown in
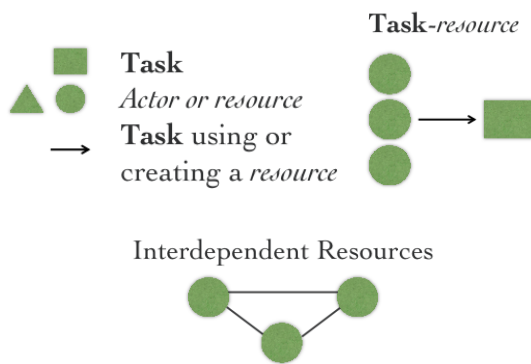

Figure 1. Tasks (squares) and resources (triangles) or actors (circles) (upper left); tasks create/use re-sources creating a task/resource assignment problem (upper right); and dependencies among interdependent resources (bottom).
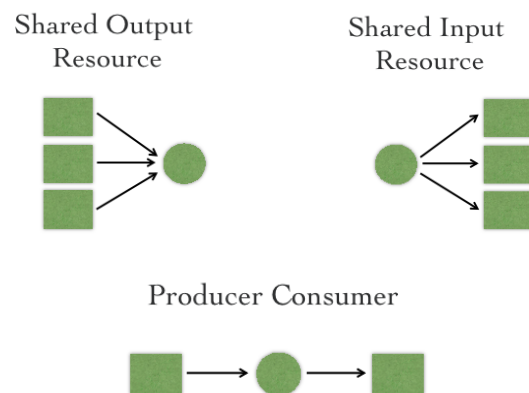
Figure 2. Three types of dependencies between two tasks (squares) based on shared use of resources (circles).

Figure 2. First, *producer-consumer* or flow dependencies match Thompson's sequential dependency (Thompson 1967): One task creates a resource that a second uses. For example, in a data analysis pipeline, the flow of data from one analysis to another creates a dependency between those tasks. Flow dependencies further imply the need to manage the *usability* of the resource and the *timing* and *location* of its availability (that is, a flow dependency has three aspects), e.g., data from one stage of an analysis pipeline must be suitable for the next stage and made available on time.

Second, a *shared-output* or fit dependence occurs when two activities collaborate in the creation of an output (in the case where the output is identical, there is potential synergy, because the duplicate work can be avoided). For example, data analyses to support a paper need to be tailored to work together.

Finally, a *shared-input* dependency emerges among activities that use a common resource (like Thompson's pooled dependency). For example, data collection might require a specific scientific instrument or the time of a human observer, constraining how data collection tasks are done and requiring some coordination, e.g., a schedule of observation times.

Resources may also be directly interdependent due to physical connections (the right side of Figure 1), e.g., a section of a paper that refers to results established in a prior section or data sets that need to be analyzed as an ensemble. In this case, tasks that use one resource may have dependencies with another task that uses an interdependent resource.

The key point in coordination theory is that *dependencies* (of all kinds) create *coordination problems* that may require additional work to manage. Malone and Crowston (1994) called this additional work *coordination mechanisms*. For example, if expertise is necessary to perform a given task (i.e., there is a *task-actor* dependency, a special case of *task-resource*), then an actor with that expertise must be identified and the task assigned to him or her. The work of identifying

an expert and maintaining a task assignment system constitutes the *coordination mechanism*. To avoid coordination problems (e.g., a task not getting done), this additional work must be performed by someone (e.g., a manager or even the expert him or herself).

Several mechanisms often can be used to manage a given dependency. For example, to manage a *usability* dependency (part of a *flow* dependency), the resource created might be tailored to the needs of the consumer (meaning that consumers must provide information about their needs to the producer) or a producer might follow a standard so the consumer knows what to expect. *Usability* dependencies are particularly salient in scientific research. Data (a particularly important resource) that are collected must be appropriate for the research question and be credible according to the standards of the field to be useful for analysis. Analysis reports (another resource) must meet the expectations of the field and provide answers to questions of interest to be useful for paper writing. Papers must be written in the genre of a scientific paper to be publishable, with the details of the genre differing from field to field. An important part of the training of a scientist is to learn the specific expectations for data, analysis reports, and papers in the scientist's research field. That is, the expertise required to accomplish a task includes knowing how to do it in the way expected by users of the output.

Note that in developing the coordination theory framework, Malone and Crowston (1994) describe coordination mechanisms as relying on other necessary group functions, such as decision making, communications, and development of shared understandings and collective sense-making (Crowston and Kammerer 1998). To develop a complete model of a process would involve modeling all these aspects. In this paper, though, we will focus on the coordination aspects, mostly bracketing the other phenomena.

In summary, coordination theory provides a lens (i.e., a set of concepts) with which to analyze group processes: In terms of tasks, resources, dependencies, and coordination mechanisms. The theory posits that the fit or lack of fit between the dependencies and available coordination mechanisms may explain problems faced by the group in achieving its goals. Finally, the framework provides a way to structure the analysis by listing possible dependencies to look for.

Coordination in Virtual Citizen Science Projects and in Collaborative Paper Writing

In this section, we present a theoretical analysis of virtual citizen science projects from a coordination-theory perspective as a basis for then analyzing the coordination challenges of the work of volunteers in Galaxy Zoo Quench. We start by presenting an analysis of the coordination challenges of work in virtual citizen science projects. We specifically describe Galaxy Zoo, which is a prototypical virtual citizen science project, as well as the basis for the Quench project, i.e., the work with which the Galaxy Zoo Quench volunteers were familiar. This analysis describes most virtual or online data-analysis citizen science projects, but not action or conservation projects (Wiggins and Crowston 2011). We then develop an analysis of the coordination challenges faced by volunteers in writing a paper, which was the ultimate goal of the Quench project.

Virtual citizen science.

A specific example of a virtual citizen science project is Galaxy Zoo (http://galaxyzoo.org/), for which volunteers support scientific inquiry by online analysis of the millions of astronomical photographs collected by the Sloan Digital Sky Survey and others. Specifically, the Galaxy Zoo system asks individuals to answer a series of questions about the shape of a galaxy captured in an image (e.g., the number of spiral arms or how round or elliptical they are) to support astronomical research on galaxy morphology. The workflow for the data-analysis task is shown in Figure 3. Our analysis is based on our own experience with Galaxy Zoo and other virtual citizen science projects

as well as published studies of these projects (e.g., Prestopnik and Crowston 2012; Wiggins and Crowston 2015; Tinati et al. 2015; Simpson et al. 2014).

One coordination problem in these projects results from a *task-actor* dependency that requires a *task assignment* coordination mechanism to match a task (an image to be classified) to an actor (a volunteer). This dependency is handled by the system simply giving the next image to be classified to the next available volunteer who has not already seen it (Reeves et al. 2017). This approach has the advantage of being simple and requiring almost no information about the image or volunteer.

A second problem is ensuring data quality, that is, the *usability* of the classifications as they flow to the research project. The quality of citizen science data is a key issue for many projects (Riesch and Potter 2014; Wiggins et al. 2011). In these projects, this usability dependency (an aspect of a flow dependency) is handled by having multiple volunteers repeat the classification and using the consensus to eliminate occasional classification errors.

Figure 3 also shows a flow of data from occasional serendipitous discoveries (the yellow dot in the figure). Every image is inspected by human analysts who may identify oddities in the images, such as the Voorwerp discovered in Galaxy Zoo. As the figure shows, such discoveries
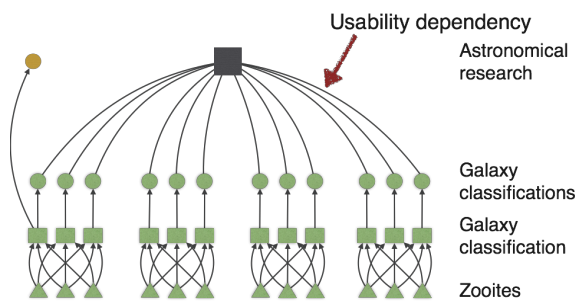

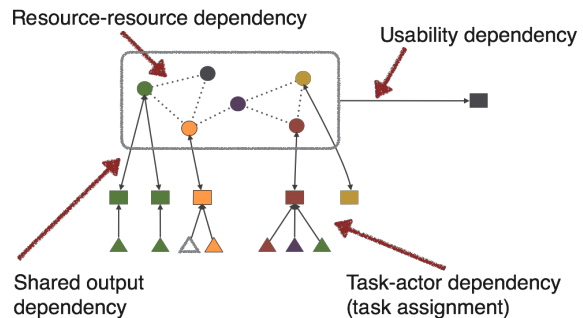**Figure 3.** Flow of data in the Galaxy Zoo project.

Figure 4. Expected structure of dependencies in writing a paper.

are handled outside the regular project flow, and they support research other than the planned project research (Tinati et al. 2015).

In summary, the main tasks of virtual citizen science projects such as Galaxy Zoo have minimal dependencies, and the coordination mechanisms needed are easily provided by the Zooniverse system.

Paper writing. In contrast to classifying data for citizen science, writing a scientific paper includes more varied dependences that potentially pose challenges to the writers. Figure 4 shows the structure of dependencies involved, based on published work on coordination in writing (Erkens et al. 2005), Wikipedia in particular (e.g., Kittur et al. 2009), and a detailed coordination-theory analysis of a comparable process, writing software (Crowston and Scozzi 2002; Crowston and Scozzi 2008). There are also differences in the diversity and difficulty of the tasks included, but these are not considered in analyzing the task dependencies.

A first difference between Figures 3 and 4 is the presence of *dependencies among the parts of the paper*, the outputs of the paper-writing tasks. Only a few tasks in writing, such as proofreading, are like galaxy classification in that they can be done without affecting other tasks (Kittur et al. 2009), e.g., by crowdsourcing (Bernstein et al. 2015). For the most part, different parts of a paper cannot be written independently. For example, the research problem presented in the introduction must be supported in the literature review, answered in the data analysis, and so on (Wichmann and Rummel 2013). Furthermore, the voice and writing style of the different sections need to match. These dependencies among paper parts impose constraints on how those parts are written (Kittur et al. 2009), posing coordination challenges to the people working on them.

To manage these dependencies requires additional work: Authors must either plan the writing process in advance (Viégas et al. 2007; Erkens et al. 2005), e.g., by developing a shared vision

for the paper (Wichmann and Rummel 2013) collectively or led by one person (Kim et al. 2014), or they must write and revise their parts to fit with other parts (i.e., mutual adjustment). Teevan et al. (2016) report on a system to create microtasks for paper writing, but despite the design intent of minimizing dependencies, they observed "considerable interaction among group members" using the system.

A second dependency is a *shared-output* dependency, created when two authors work on tasks that have the same output, i.e., two authors working on writing the same part of the paper. Galaxy Zoo also has multiple volunteers work on the same galaxy image, but because the possible results are few in number, a simple consensus rule is usually sufficient to merge the classifications. However, many more differences can arise in writing a paper. At a basic level, problems of simultaneous changes to text can be managed by a shared document editor (Lowry and Nunamaker 2003). However, problems can arise at a conceptual level that are more difficult to identify and resolve (Erkens et al. 2005). To manage this dependency requires a technique to mitigate these possible conflicts in output, e.g., picking one version and rejecting the others or manually merging the changes.

A third dependency is the *task-actor* dependency. Unlike the system assignment in Galaxy Zoo, volunteers working on a paper will likely choose for themselves which tasks to work on, as they do in Wikipedia. Reliance on self-assignment of tasks fits the voluntary nature of the project but raises two potential problems.

First, people choosing to work on one part of the paper may not be good at writing, i.e., their contributions might not be usable. In a conventional team, members would be assigned to tasks based on skills, but in a voluntary setting, skills are not guaranteed. A paper-writing process must include mechanisms to assess if a writing contribution is acceptable (Kittur et al. 2007). For

example, in Wikipedia, editors police edits and modify or revert problematic ones. As an alternative approach, rather than rejecting substandard results, efforts could be made to provide the volunteer with the skills necessary to achieve a good outcome, e.g., by providing training.

Second, a volunteer might not be reliable, meaning that a promised contribution might not appear on time (Riesch and Potter 2014). The writing process will thus need mechanisms to handle missing contributions. This problem interacts with the second dependency, shared output, because one way to minimize problems from the former issue is to have only one person at a time work on a task (i.e., assign authors for each document section). Such a process is problematic, however, if there is a chance that the task (i.e., the document section) will not be completed.

A final dependency is a *usability* dependency (part of the *flow* dependency) between the creation of the paper and the use of the paper by its intended audience. In the basic work of citizen science projects, the usability of the resulting data set is managed by having the science teams design the process of creating the data, with carefully imposed quality checks (Simpson et al. 2014). For scientific writing, this dependency is handled in part by processes such as peer review that check for article quality. However, much of the process is handled by the authors themselves acting as proxies for the readers. Knowing the scientific literature, scientific authors pick topics and write in ways that they know will be useful for that community (e.g., in the genre of a scientific article). A volunteer-driven writing process will need ways to provide information about the needs and desires of the readers to the volunteer authors, who again cannot be assumed to have specific knowledge needed to make these decisions themselves.

In summary, the task of writing a paper displays more and more varied dependencies than a prototypical citizen science project, thus posing additional coordination challenges to the volunteers. As a result, in the Quench project, we expected to see either additional work done to manage

these dependencies, or problems arising from these dependencies going unmanaged. Identifying the kinds of coordination mechanisms created or needed will be informative for managers of citizen science projects interested in involving volunteers in these additional kinds of scientific work and by extension, to other crowd researchers.

Methods

Our study of the coordination challenges faced by volunteers working on the Zooniverse Quench is a multi-method study that combines collaborative basic research (van de Ven 2007) and coordination analysis (Crowston and Osborn 2003) to understand the design and outcomes of an example of a specific kind of online citizen science. By design, we mean the tasks that the volunteers were asked to work on. By outcomes, we mean the results of those tasks. We worked in close collaboration with developers, designers, and educators at Zooniverse.

Data elicitation included questionnaires, interviews, and focus groups addressing volunteer motivation and learning as well as trace data analysis on volunteers' experiences with the project. There were two rounds of questionnaires, one sent before and one after the project to a mailing list with 1697 members. The questionnaires included both closed-end (agree/disagree or Likert scale) questions and open-ended questions. The pre-survey, which received 490 responses (a 29% response rate), asked about activities in other Zooniverse projects, feeling of participation in scientific process, perceived value of contribution to project, and perceived scientific skills (e.g., making a plot). The post-survey asked about participation in the different stages of the project, reasons for participating, activities that were enjoyable or that were challenging, valuable or non-valuable resources, perceived scientific skills, improvement in skills from the project, enjoyment of the project, and demographics. The post-survey received 183 responses (an 11% response rate) of which 74 were complete; 60 of the post-survey respondents had completed the pre-survey.

Responses to the open-ended questions were inductively content analyzed to identify a set of themes regarding volunteers' perceptions of the project and the analysis work. Open coding generated many common themes, which were grouped into nine themes exploring users' perceptions of the project: Preference for classification, unclear guidance, data problems, tools, lack of time, lack of priority, lack of confidence in abilities, lack of access to resources, and lack of awareness. These themes are discussed below.

In addition, we conducted extensive analysis of the discussion board associated with the project. The goal of this analysis was to map the history of the project and important events and decisions made over the course of the project. A list of talk posts referenced in the paper is given in Table 1. Finally, the study draws on several years of the authors' prior engagement with the broader citizen science community.

Table 1. Galaxy Zoo Quench talk posts referenced in the paper.

ID	Title	URL
a	Galaxy Zoo Quench Project Overview	https://quench.galaxyzoo.org/#/project
b	Major redshift measurement errors in the SDSS stereoscopic pipeline	BGS000000b/discussions/DGS000021u
c	Quench Talk Office Hours	BGS000000a/discussions/DGS00001xk
d	Sample Selection: Post-quenched galaxy and control galaxy	BGS0000001/discussions/DGS00001xy
e	Classification Result Error	BGS0000008/discussions/DGS000020s
f	Difference between v5 QS and QC catalogs and their v4 counterparts	BGS0000008/discussions/DGS000022a
g	Dealing with Sample Selection Issues	BGS0000008/discussions/DGS0000223
h	Framework and Suggestions for Data Analysis Phase	BGS0000007/discussions/DGS000013u
i	Temporary Pause in Quench (Resume July 25th)	BGS000000f/discussions/DGS000023b
j	Quench project: a proposal aimed at reviving and completing it	BGS000000e/discussions/DGS000022f

[1] Unless a complete URL is given, URLs start https://quenchtalk.galaxyzoo.org/#/boards/

*C*oordination analysis Crowston and Osborn (2003) led us to pay attention to dependencies in the tasks in Galaxy Zoo Quench. The analysis has six steps: Defining process boundaries, collecting data, determining actors and resources, determining activities, determining dependencies, and model verification. We analyzed our data using this technique, which highlighted dependencies in the system and led to our documentation of the coordination process associated with each project. Equally important, this technique reveals areas where the management of dependencies broke down, causing coordination problems.

Results: Coordination Problems in Advanced Citizen Science Work

We turn next to an examination of the dependencies, coordination mechanism, and observed coordination problems in Galaxy Zoo Quench. Figure 5 presents the planned flow of data throughout the project, indicating in grey boxes the major outcomes of each of the phases. The plan was to complete Phase 1, the classification process, by 1 August 2013 and then proceed to Phase 2, data analysis and discussion. The goal for Phase 3 was for the volunteers (working with one of the Galaxy Zoo scientists) to write and submit a 4–5 page article to *Monthly Notices of the Royal Astronomical Society (MNRAS) Letters*, the online portion of the *MNRAS Journal*. The color of the arrows indicates the success of each phase, with green indicating success, yellow indicating partial success, and red indicating incompletion, as assessed by an analysis of the discussion around each phase. We next describe each of the phases in more detail with attention to the coordination challenges encountered.

Phase 1. Phase 1 of the Quench project focused on the tasks of classifying galaxies. In this phase 1, the volunteers coded galaxies to find those with the characteristics of post-quenched galaxies. The classification included characteristics believed to be related to quenching, specifically
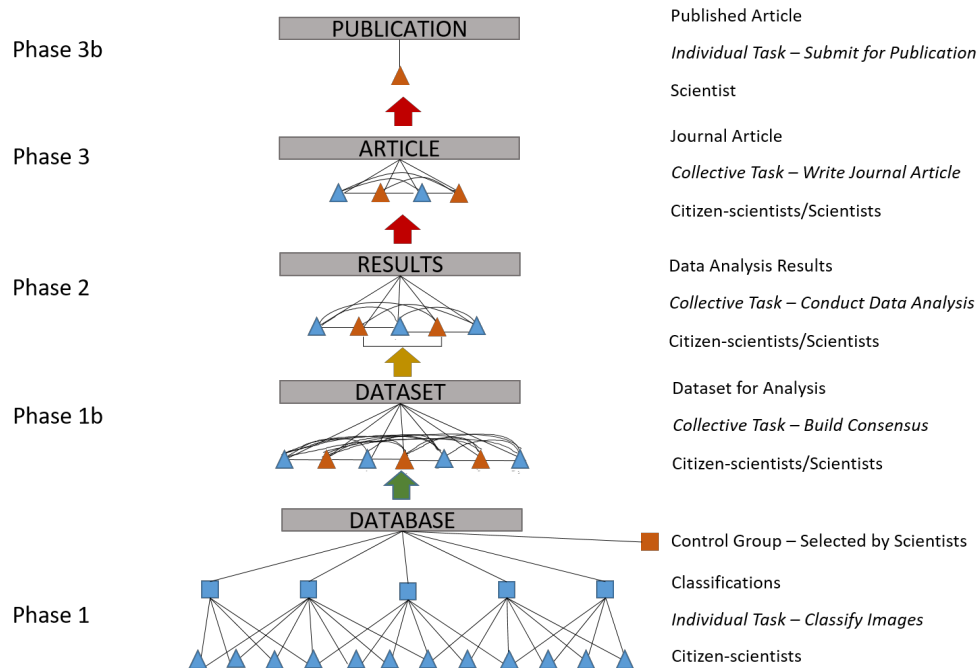
Figure 5. Expected structure of dependencies in Galaxy Zoo Quench
(data flows from the bottom up).

galaxies merging, tidal debris, both, or neither. Galaxy classification is a mature process that has been used on many citizen science projects, mostly notably Galaxy Zoo. The classification was somewhat delayed, but successfully completed in that the selected galaxies had classifications by the end of August 2013 (Source a). Of the respondents to the post-survey, 129 (70%) reported working on Phase 1.

Phase 1b. The project as executed included an additional phase between the initial coding and data analysis, labelled Phase 1b in Figure 5. This phase, which was not explicitly described in the announced plans for the Galaxy Zoo Quench project, represents the first collective task in the process, building consensus on the data created by the individual classifications made by the volunteers in Phase 1 to generate a dataset for analysis in Phase 2. The assumption was that once the galaxies were coded, the results could be used for analysis. In fact, reaching consensus on the dataset turned out to be a significant undertaking for the group.

16

First, as the volunteers started to use the data, they raised concerns about how the final classification was assigned (Sources b, c & d). The initial algorithm used to determine a classification was to take the option selected by the most volunteers, as in other Galaxy Zoo projects (Source e). For example, if "merging" was selected by 3, "tidal debris" by 6, "both" by 2, and "neither" by 7, the galaxy would be classified as "neither," even though together the other choices, which indicate an interesting finding, had been chosen more often. This discrepancy was fixed by revising the algorithm to add the count of the three interesting findings together.

Second, the process of revising the dataset led to concerns about the usability of the data. Two of the themes in the post-survey were data and tool problems. One respondent commented that "data sets were slow to become available, they contained errors and the analysis tools were immature or buggy." For instance, volunteers were uncertain about the data reliability given the significant changes made between versions. In some cases, errors crept into the files as they were processed by different people. For example, identifiers for the galaxies in the data file are 18-digit numbers. If the file is opened in Excel (a common tool for citizen scientists and scientists more generally because it is widely available), these long numbers could be converted to floating point numbers and truncated, changing the ID, a problem that beset some versions of the data file. There was also inconsistency in variable labeling between datasets, which raised questions about the data provenance (Source f).

A third set of questions arose about the control group of galaxies. To provide a comparison to the quenched galaxies, the scientists involved in the project selected a control group of 3,002 galaxies but did so independently from the volunteers. The citizen scientists requested clarification on the selection of the control group, which was explained, but doubts remained (Source g). Throughout the project, a recurrent discussion involves the suitability of the sample of galaxies for

the study. Sampling had to be done carefully to avoid introducing bias into the results. Participants developed different subsamples based on different selection rules, but did not seem to reach consensus about which sample should be used.

Phase 2. The next phase of the project, Phase 2 in Figure 2, was data analysis. The lead scientist working with the volunteers had encouraged them to "play" with the data and to "have fun and ferret out interesting trends in the data" (Source h). The intent was that the volunteers would explore on their own and then share interesting results with the group, thus experiencing the process of scientific discovery.

As noted above, volunteers had already been observed engaging in analyses of other data sets in other projects. While about half of the respondents to the post-survey commented on a perceived lack of confidence in their own abilities (or a lack of time that would be needed to gain the skills), others seemed capable of such work. Figure 6 shows the responses to questions on the pre-survey about data analysis skills, indicating that the majority of respondents felt that they had relevant skills (with the exception of writing a paper). The subset of volunteers who worked on Phase 2 seemed even more knowledgeable. Indeed, a number of the responses to the post-survey note the high level of the discussion as a deterrent to participation, e.g., "The knowledge and ability of the remaining participants so far exceeded my own knowledge." Furthermore, different analyses could be done in parallel, i.e., there was no dependency between the analyses that would require the volunteers to coordinate their work on the analyses.

Unexpectedly, though, the group encountered difficulties in this phase. Volunteers perceived the task as too open-ended and so did not know how best to proceed. A theme in the post-survey was unclear guidance. For example, one respondent commented that he "was unsure how to report my results, and on a more basic level was unsure about which descriptive and inferential
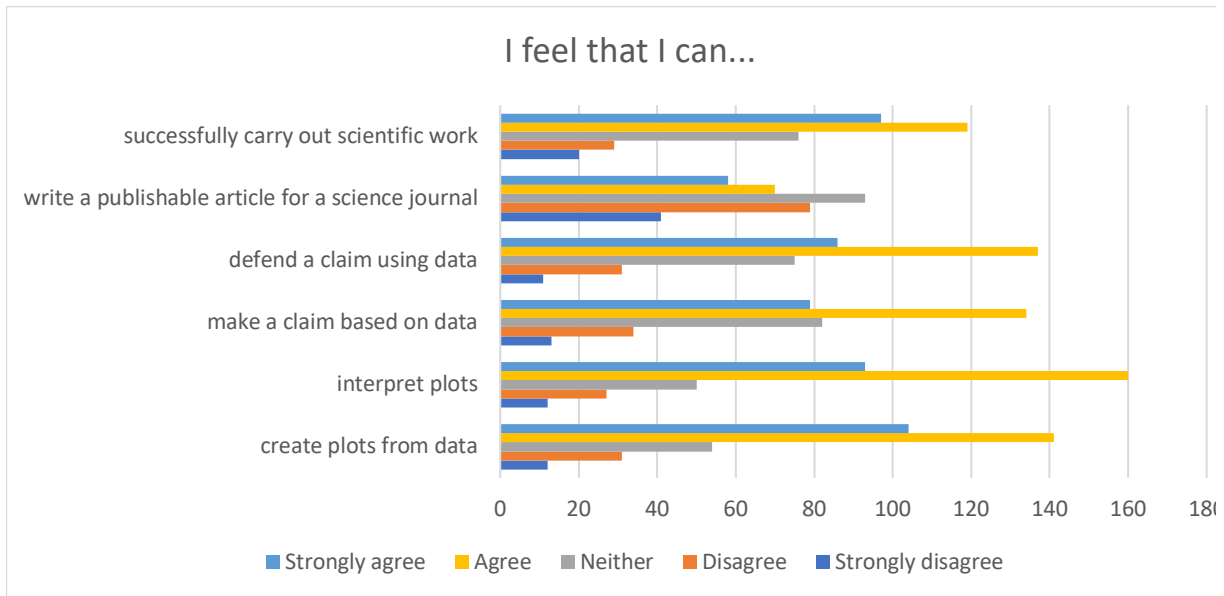
Figure 6. Histogram of responses to questions regarding perceived data analysis and reporting skills (N = 490)

stats were the most relevant to the project." Another stated more bluntly, "I never could figure out what they wanted, what the point was, so, I was there for maybe 20 minutes, clicked out, VERY confused, and never went back." Part of the volunteer feedback on the project was that the project needed more scaffolding of the research process.

Unfortunately, during this phase, the lead scientist became unavailable for some time and none of the other scientists on the project could take on a leadership role (Source i; also noted on the post-survey). Problems caused by the absence of a single key individual would not be surprising in a conventional team, but they were unexpected in the context of a citizen science project in which members were able—and expected—to make independent contributions. The volunteers attempted to continue the project, with extensive discussion and various analyses developed. However, the volunteers did not reach a final decision about what should be done, so Phase 2 did not progress to having the desired final set of analyses and a scientific story. As a final analysis was not completed, Phase 3, writing, never started.

19

In 2014 (and again more recently), a citizen scientist attempted to revive the project, receiving responses from the other citizen scientists, as well as from three scientists (Source j). However, the discussion ended without the project restarting, and no further posts appeared on the Galaxy Zoo Quench Talk board.

Discussion

In this section, we interpret the case using coordination theory to identify what kinds of dependencies existed, how those dependencies were managed or not managed, and the impact of these dependencies on project performance.

Phase 1 of the Quench project focused on the tasks of classifying galaxies. Participants could work independently and concurrently to classify the post-quenched galaxies, with minimal dependencies creating constraints on their work. Classification is a mature process, with a sound technological platform and significant history of being completed in Galaxy Zoo, as well as other citizen science projects. The task of looking at an image and clicking on classifications is well defined. A number of the post-survey responses note familiarity with this process, e.g., "As a long time zooite I felt very comfortable with the classification stage." Indeed, a theme in the post-survey was that some participants preferred classification to the other work. Citizen scientists were both producers and consumers of the data, at least for those continuing to participate beyond Phase 1, so they had significant motivation to complete the task in a timely fashion. As a result, Phase 1 was completed successfully.

In Phase 1b, the volunteers undertook several tasks to refine the data set for analysis. There is a dependency among these various data refinement tasks because they are contributing toward a common output. For example, an important part of the analysis was determining which galaxies to include or to exclude in a way that did not introduce biases in the sample that would affect the

20

results. In this phase, the volunteers started to experience difficulties ensuring that the decisions were made consistently. Another interpretation is that the analysis task has a usability dependency with the creation of the data set, and the various steps undertaken to refine the data are ways to manage this dependency. However, it was difficult for the volunteers to know what criteria were appropriate, given their lack of expertise in this form of analysis.

Phase 2 seemed to suffer from more significant coordination problems. First, in an effort to encourage exploration and serendipitous discoveries, the project scientists provided only general guidance about what analyses should be done, planning to react to the findings of the volunteers. The problem experienced by the volunteers was parallel to the difficulties in developing a suitable data set: Even when they have the skills to do an analysis, the volunteers do not have the expertise to know which analyses will be suitable for publication, so they cannot ensure the usability of their output for the next phase, paper writing. One respondent to the post-survey suggested that "there does need to be some comments by the science team as to whether what is being discussed is accurate and/or relevant."

Our initial expectation was that the project would face challenges particularly in Phase 3, owing to the number of dependencies that would have to be managed to generate a coherent paper. However, as Phase 3 did not start, this case does not provide data on this question.

In summary, our analysis of the dependencies in the project suggest that a key problem throughout was ensuring the usability of the outputs of each phase of the project for the next phase. In the first phase, the usability of the galaxy classifications was ensured by the design of the coding system and of the Zooniverse system. Even here, issues arose because the coding system was more complicated, requiring a different aggregation technique. Next, creating a data set that was suitable for analysis (Phase 1b) required not only coding galaxies but also selecting a suitable sample,

which requires expertise to do in an acceptable way. Finally, in Phase 2 the project ran into unexpected difficulties in finalizing a set of analysis results that would support a paper. Because the volunteers were not experts in astrophysics, it did not seem possible for them to say what analyses would be suitable. It is interesting to speculate what would have been the result in the Quench case if the volunteers had been given more specific direction on which analyses to run. However, this approach would have in some ways been contrary to the project goal to allow volunteers to engage in discovery on their own.

Another way to express the problem experienced is that there was a need to decompose the overall task of developing an analysis into more specific subtasks that different volunteers can work on. However, the volunteers lacked the knowledge of astrophysics (for Quench) needed to do this decomposition and the concomitant recombination. As a result, when the project scientists were unable to give guidance, the analysis process stalled, with volunteers unsure what actions to take.

The continued interest of the volunteers suggests that the project did not suffer from a lack of motivation on their part. Nor does expertise in data analysis seem to have been the issue: At least some participants felt confident to go on, judging by the responses to the survey and the comments of participants. However, the task of managing the usability dependency between analysis and paper writing (and to some extent, between data collection and analysis) turned out to be one that could not be entirely delegated to a volunteer, no matter how motivated.

Conclusions

From our initial analysis, we expected that citizen scientists involved in Galaxy Zoo Quench would encounter problems coordinating the work of writing a paper due to the increased coordination demands of this task as compared to the low level of dependencies in typical citizen science work.

Unexpectedly, the Quench project encountered significant difficulties at the prior phase of developing a dataset and had difficulties conducting analyses, even though volunteers had an interest, motivation, and prior demonstrated ability to conduct analyses, and in principle the tasks to be done had few dependencies, as different analyses could be carried out separately.

This case thus provides a theoretical perspective on the strengths and limitations of the virtual citizen science model. Virtual citizen science has been strikingly successful in tapping the interest, capabilities, and energies of a large number of volunteers to carry out analyses of datasets that are beyond the capacity of science teams. However, giving such work "to the crowd" is feasible when there are minimal dependencies between tasks (e.g., a pooled dependency of determining a consensus) and where the desired output can be clearly described. Extending citizen science work to tasks with more dependencies or unspecified outputs strains this model.

Our analysis leads to several recommendations for how to support advanced work with citizen scientists. First, it is important to have a complete accounting of what the relevant tasks are. The analysis done in designing the Quench project seems to have overlooked the work that precedes and surrounds specific analyses. Specifically, the scientists did not seem to account for the work that must be done to ensure that a dataset is usable for analysis or to select which analyses will be interesting to perform. It may be that for experienced researchers, this type of work "goes without saying," but in a citizen science setting, it needed to be spelled out. Furthermore, given the reliance of citizen science on volunteers self-selecting tasks, it is necessary to carefully analyze the tasks to ensure that they are feasible for volunteers. Coordination theory can be useful for examining the dependencies between tasks that might constrain the work.

The main issue highlighted by this case is the difference between knowing how to do a task and knowing what users of the output will find useful. It seems that much of the work of ensuring

the usability of outputs required tacit knowledge, in this case about what data should look like or what analyses are interesting for publication. The problem of volunteers evaluating their results has been noted in other crowdsourcing settings (Nagar et al. 2016) and for open innovation projects (Schulze et al. 2012). For a task to be suitable for crowd work, evaluation criteria need to be made explicit. Østerlund and Crowston (in press) have similarly noted the need for more explicit instructions to enable communication between groups that do not share common background knowledge.

Citizen science projects sometimes provide training, which can be quite intensive (e.g., in the details of a data collection protocol). Training can address the concern noted above of a volunteer's being able to do a task. However, it does not seem feasible to train volunteers to develop the insight needed to know what kinds of data or analyses will be interesting for publication. Indeed, even advanced graduate students in a topic can struggle with these questions. Instead, we recommend that projects faced with these sorts of usability dependencies implement feedback mechanisms to quickly evaluate proposals from the volunteers and to provide guidance on improving them, likely the original plan for Galaxy Zoo Quench. However, it is hard to know whether feedback alone would be enough to guide volunteers to a publishable or scientifically useful analysis result.

Finally, our analysis of the coordination needed for collaborative writing suggests that the volunteers would have faced significant challenges had they gotten to Phase 3. Exploring the kinds of challenges involved in this sort of work remains a topic for further research.

Acknowledgements

References

Bernstein, M. S., Little, G., Miller, R. C., Hartmann, B., Ackerman, M. S., Karger, D. R., Crowell, D. & Panovich, K. 2015. Soylent: A word processor with a crowd inside. *Communications of the ACM,* 58**,** 85-94.

Bonney, R., Cooper, C. B., Dickinson, J., Kelling, S., Phillips, T., Rosenberg, K. V. & Shirk, J. 2009. Citizen science: A developing tool for expanding science knowledge and scientific literacy. *BioScience,* 59**,** 977–984.

Crowston, K. 1997. A coordination theory approach to organizational process design. *Organization Science,* 8**,** 157–175.

Crowston, K. & Kammerer, E. 1998. Coordination and collective mind in software requirements development. *IBM Systems Journal,* 37**,** 227–245.

Crowston, K. & Osborn, C. S. 2003. A coordination theory approach to process description and redesign. *In:* Malone, T. W., Crowston, K. & Herman, G. (eds.) *Organizing Business Knowledge: The MIT Process Handbook.* Cambridge, MA: MIT Press.

Crowston, K. & Scozzi, B. 2002. Open source software projects as virtual organizations: Competency rallying for software development. *IEE Proceedings Software,* 149**,** 3–17.

Crowston, K. & Scozzi, B. 2008. Bug fixing practices within Free/Libre Open Source Software development teams. *Journal of Database Management,* 19**,** 1–30.

Erkens, G., Jaspers, J., Prangsma, M. & Kanselaar, G. 2005. Coordination processes in computer supported collaborative writing. *Computers in Human Behavior,* 21**,** 463–486.

Galbraith, J. R. 1973. *Designing Complex Organizations,* Reading, MA, Addison-Wesley.

Kaplan, A. 1964/1998. *The Conduct of Inquiry: Methodology for Behavioural Science,* New Brunswick, NJ, Transaction Publishers.

Kim, J., Cheng, J. & Bernstein, M. S. Ensemble: Exploring complementary strengths of leaders and crowds in creative collaboration. In: Proceedings of the ACM Conference on Computer Supported Cooperative Work & Social Computing, Baltimore, Maryland, USA, pp. 745–755.

Kittur, A., Lee, B. & Kraut, R. E. Coordination in collective intelligence: The role of team structure and task interdependence. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Boston, MA, USA, pp. 1495-1504.

Kittur, A., Suh, B., Pendleton, B. A. & Chi, E. H. He says, she says: Conflict and coordination in Wikipedia. In: Conference on Human Factors in Computing Systems, San Jose, California, USA, pp. 453-462.

Lintott, C. J., Schawinski, K., Keel, W., Van Arkel, H., Bennert, N., Edmondson, E., Thomas, D., Smith, D. J. B., Herbert, P. D., Jarvis, M. J., Virani, S., Andreescu, D., Bamford, S. P., Land, K., Murray, P., Nichol, R. C., Raddick, M. J., Slosar, A., Szalay, A. & Vandenberg, J. 2009. Galaxy Zoo: 'Hanny's Voorwerp', a quasar light echo? *Monthly Notices of the Royal Astronomical Society,* 399**,** 129-140.

Lowry, P. B. & Nunamaker, J. F. 2003. Using Internet-based, distributed collaborative writing tools to improve coordination and group awareness in writing teams. *IEEE Transactions on Professional Communication,* 46**,** 277-297.

Malone, T. W. & Crowston, K. 1994. The interdisciplinary study of coordination. *Computing Surveys,* 26**,** 87–119.

Nagar, Y., De Boer, P. & Bicharra Garcia, A. C. Accelerating the review of complex intellectual artifacts in crowdsourced innovation challenges. In: Proceedings of the International Conference on Information Systems, Dublin.

Newman, G., Wiggins, A., Crall, A., Graham, E., Newman, S. & Crowston, K. 2012. The future of citizen science: Emerging technologies and shifting paradigms. *Frontiers in Ecology and the Environment,* 10**,** 298–304.

Østerlund, C. & Crowston, K. in press. Documentation and access to knowledge in online communities: Know your audience and write appropriately? *Journal of the Association for Information Science and Technology*.

Owens, T. 2016. Making crowdsourcing compatible with the missions and values of cultural heritage organisations. *Crowdsourcing Our Cultural Heritage***,** 269.

Prestopnik, N. R. & Crowston, K. Citizen science system assemblages: Understanding the technologies that support crowdsourced science. In: iConference, 7–10 February, Toronto, Ontario.

Raddick, M. J., Bracey, G., Gay, P. L., Lintott, C. J., Murray, P., Schawinski, K., Szalay, A. S. & Vandenberg, J. 2010. Galaxy Zoo: Exploring the motivations of citizen science volunteers. *Astronomy Education Review,* 9**,** 010103–18.

Reeves, N., Tinati, R., Zerr, S., Kleek, M. G. V. & Simperl, E. From crowd to community: A survey of online community features in citizen science projects. In: Proceedings of the Conference on Computer Supported Cooperative Work and Social Computing, Portland, Oregon, USA, pp. 2137-2152.

Riesch, H. & Potter, C. 2014. Citizen science as seen by scientists: Methodological, epistemological and ethical dimensions. *Public Understanding of Science,* 23**,** 107-120.

Rotman, D., Hammock, J., Preece, J., Hansen, D., Boston, C., Bowser, A. & He, Y. Motivations affecting initial and long-term participation in citizen science projects in three countries. In: iConference 2014 Proceedings, pp. 110–124.

Rotman, D., Preece, J., Hammock, J., Procita, K., Hansen, D., Parr, C., Lewis, D. & Jacobs, D. Dynamic changes in motivation in collaborative citizen-science projects. In: Conference on Computer Supported Cooperative Work, Seattle, Washington, USA, pp. 217-226.

Schulze, T., Indulska, M., Geiger, D. & Korthaus, A. Idea assessment in open innovation: A state of practice. In: 20th European Conference on Information Systems, 11-13 June, Barcelona, Spain, pp. Paper 149.

Simpson, R., Page, K. R. & Roure, D. D. Zooniverse: Observing the world's largest citizen science platform. In: Proceedings of the 23rd International Conference on World Wide Web, Seoul, Korea, pp. 1049-1054.

Teevan, J., Iqbal, S. T. & Veh, C. V. Supporting collaborative writing with microtasks. In: Proceedings of the CHI Conference on Human Factors in Computing Systems, Santa Clara, California, USA, pp. 2657–2668.

Thompson, J. D. 1967. *Organizations in Action: Social Science Bases of Administrative Theory,* New York, McGraw-Hill.

Tinati, R., Kleek, M. V., Simperl, E., Luczak-Roesch, M., Simpson, R. & Shadbolt, N. Designing for citizen data analysis: A cross-sectional case study of a multi-domain citizen science platform. In: Proceedings of the Conference on Human Factors in Computing Systems, Seoul, Republic of Korea, pp. 4069-4078.

Van De Ven, A. H. 2007. *Engaged Scholarship: A Guide for Organizational and Social Research,* Oxford, UK, Oxford University

Viégas, F. B., Wattenberg, M., Kriss, J. & Van Ham, F. Talk before you type: Coordination in Wikipedia. In: Hawai'i International Conference on System Sciences (HICSS-40).

Wichmann, A. & Rummel, N. 2013. Improving revision in wiki-based writing: Coordination pays off. *Computers & Education,* 62**,** 262-270.

Wiggins, A. & Crowston, K. From conservation to crowdsourcing: A typology of citizen science. In: Hawai'i International Conference on System Science (HICSS-44), January, Koloa, HI.

Wiggins, A. & Crowston, K. 2015. Surveying the citizen science landscape. *First Monday,* 26.

Wiggins, A., Newman, G., Stevenson, R. D. & Crowston, K. Mechanisms for data quality and validation in citizen science. In: Proceedings of the International Conference on e-Science, pp. 14-19.

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