When Scientists Meet the Public: An Investigation into Citizen Cyberscience

Peter Darch

Corpus Christi College

University of Oxford

A dissertation submitted in partial fulfilment

of the requirements for the degree of

Doctor of Philosophy

Michaelmas Term 2011

Peter Darch
Corpus Christi College
When Scientists Meet the Public: An Investigation into Citizen Cyberscience

Abstract

Citizen Cyberscience Projects (CCPs) are projects mediated through the Internet, in which teams of scientists recruit members of the public (volunteers) to assist in scientific research, typically through the processing of large quantities of data. This thesis presents qualitative ethnographic case studies of the communities that have formed around two such projects, climateprediction.net and Galaxy Zoo. By considering these social actors in the broader contexts in which they are situated (historical, institutional, social, scientific), I discuss the co-shaping of the interests of these actors, the nature of the relationships amongst these actors, and the infrastructure of the projects and the purposes and nature of the scientific work performed.

The thesis focusses on two relationships in particular. The first is that between scientists and volunteers, finding that, although scientists in both projects are concerned with treating volunteers with respect, there are nevertheless considerable differences between the projects. These are related to a number of interconnecting factors, including the particular contexts in which each project is embedded, the nature of the scientific work that volunteers are asked to undertake, the possibilities and challenges for the future development of the projects as perceived by the scientists, and the tools at the disposal of the respective teams of scientists for mediating relationships with volunteers. The second is amongst the volunteers themselves. This thesis argues that volunteers are heterogeneous, from disparate backgrounds, and that they sustain their involvement in CCPs for very different purposes. In particular, they seek to pursue these through the way they negotiate and construct their relationships with other volunteers, drawing on particular features of the project to do so.

This thesis contributes to two fields. The first is to Citizen Cyberscience itself, with a view to improving the running of such projects. Some social studies have already been conducted of CCPs to this end, and this thesis both extends the analysis of some of these pre-existing studies and also problematizes aspects of CCPs that these studies had not considered. I discuss the significance of my findings for those involved in setting up and running a CCP, and present some recommendations for practice. The second field is Science and Technology Studies, in particular studies of public engagement with scientific and technological decision- and knowledge-making processes. The modes of engagement found in CCPs differ in key ways from those that have already been documented in the existing literature (in particular, different power relationships) and thus offer new ways of understanding how the public might be engaged successfully in such processes.

Acknowledgements

The completion of this thesis would not have been possible without the help and support of a number of people, to whom I am immensely grateful.

I would like to thank all the volunteers, scientists and software engineers involved in my two case studies who took the time to speak with me about their experiences in the project. In particular I would like to thank Dr Milo Thurston and Dr Tolu Aina (in the case of climateprediction.net) and Dr Chris Lintott (in the case of Galaxy Zoo) for being my first points of contact in each project, and for their useful advice. In the case of project volunteers, mo.v (in the case of climateprediction.net) and Alice Sheppard (in the case of Galaxy Zoo) gave me invaluable support for how to approach the volunteer communities in the respective projects and for helping with recruiting participants for questionnaires and interviews.

I would also like to thank deeply my supervisors, Dr Annamaria Carusi and Dr Marina Jirotka, for their advice, encouragement, support, and feedback, as well as my College advisor, Professor Colin McDiarmid, for support going back to before my undergraduate days.

Finally, I would like to thank my family and friends for their encouragement and love as I embarked on my doctoral studies, as well as for all the years before then. Above all, this thesis is dedicated to my Mum and Dad, who have given me so much support (both emotional and material), more than I could ever repay.

Table of Contents

List of Figures

Chapter 1: Introduction

In recent decades, there has been an explosion of scientific sub-disciplines and projects which involve the processing of large quantities of data or the running of computationally-intensive mathematical models (Hey *et al*. 2009; Welsh *et al.* 2006). Another critical challenge that has recently faced many scientists is a political culture, both at the national level and international level (for instance, the European Union), which increasingly calls on scientists to show a willingness to interact with and educate the lay public about their work (House of Lords 2000; Wynne & Felt 2007).

In order to meet one or both of these two significant challenges, some groups of scientists have set up *Citizen Cyberscience Projects* (henceforth, CCPs) (Haklay 2010)[1]. In these, members of the public are asked for help in conducting scientific research, and the main interface for interaction between a project's scientists and its public volunteers is via the Internet. Typically, a project is run by a team of scientists and software engineers at a particular institution (Christensen *et al.* 2005). In the light of the challenges outlined above, CCPs tend to have one or two of the following goals:

- To produce scientific results that can be disseminated; and
- To engage and educate members of the public about their field of research.

This thesis presents work consisting of qualitative sociological case studies of two CCPs, namely climateprediction.net and Galaxy Zoo. This thesis has two goals. The first is to contribute to Computer Science by using the empirical research to formulate recommendations for practice for project team members who are embarking upon CCPs or who are involved in existing CCPs. The second is to make a substantive contribution to the field of Science and Technology Studies (STS), in particular through advancing understandings of how scientists interact with members of the public, and how members of the public interact with each other, on issues of technoscience (for instance, around questions of governance of science and novel technologies).

The remainder of this Chapter describes in more detail the motivation for this research in terms of improving the efficacy of CCPs. First, I explain how it should be of interest to people involved in setting up and running such projects (Section 1.1). Then, I outline how this research should contribute to the field of STS (Section 1.2). Section 1.3 introduces my case studies and describes why they were selected. Finally, this introductory Chapter concludes with descriptions of the contents of each Chapter (Section 1.4).

[1] Citizen Cyberscience is one of a number of names that has been given to this phenomenon. Other terms include *volunteer computing* (Anderson & Fedak 2006), *public distributed computing*, and *desktop grid computing*,

1.1 Motivation for my research I: Contributions to CCPs

This Section outlines the motivation for this thesis. First, it presents an overview of the current state of CCPs. It then briefly considers work that is being conducted to improve the efficacy of CCPs in order to demonstrate how this thesis might be of interest to scientists and software engineers involved in setting up and running CCPs. With the aim of advancing understandings of how to improve the efficacy of these, this Section concludes by formulating three research questions which guide the research for the doctoral thesis.

1.1.1 Citizen Cyberscience Projects

Historic involvement of citizens in science Lay people have had a long involvement with the production and circulation of scientific knowledge, playing a variety of roles in the history of science (Vetter 2011). Indeed, up until the end of the 19th century, major novel scientific work in a range of scientific fields was conducted by many individuals who would not today be recognised as professional scientists. For instance, in the field of astronomy, major discoveries of novel astronomical objects during the 19th century were made by amateurs as well as professionals (Lankford 1981), whilst Brenna (2011) discusses the contributions of the clergy in the 18th century to botany and natural history.

Professionalization of science occurred during the 19th and early 20th century, with the rapid growth of a class of professionals who were able to make a living from scientific work (for instance, see Barton 2003; Moseley 1977). This class became differentiated from amateurs as a result of formalized university training (including the development of degree courses with specialized curricula), large-scale funding to build laboratories and procure equipment, the development of professional societies, and the situating of science within universities and new institutions. Nevertheless, in a range of scientific disciplines, lay members of the public continued to play an important, if subordinate, role in the production of scientific knowledge. Most notably, lay people have been enlisted to provide observational data that would be otherwise near-impossible, or prohibitively costly, to obtain by other methods. One well-known example of this is the annual Audubon Christmas Bird Survey which has been running since 1900, asking members of the public to count the number of each species of bird they observe in their gardens[2].

Since the advent of the Internet, many such pre-existing projects have moved online and new ones have started, finding that the Internet has helped to improve the efficacy of these projects. For instance, it is now easier for volunteers to be recruited and coordinated, and for volunteers to send data they have collected to the scientists.

---

[2] http://birds.audubon.org/christmas-bird-count (accessed 9 December 2011).

It also allows for the integration of volunteers' observational technologies with scientists' infrastructure (for instance, in the Citizen Weather Observer Program[3], volunteers are able to build home weather stations which feeds back real-time data to the scientists).

New forms of citizen science The growth of the Internet, however, has not just seen the continuation and development of citizen science projects that involve the collection of data by members of the public. Rather, new forms of public engagement in scientific projects have emerged as well, in particular projects which ask members of the public to assist in data processing and data analysis. It is this new class of projects upon which this thesis focuses. These particular CCPs can be grouped into two broad categories, depending on the extent to which volunteers play an active or passive role. Each of these categories is described in turn below.

The first category: Volunteers playing a passive role In projects in this first category, volunteers play a passive role. In these, the scientists divide the data they need to be processed into very small discrete units (known as *work units*), which are then downloaded by project volunteers onto their personal computers. These work units are processed using spare capacity on the computers and, upon completion, they are sent back to the scientists. Since the first such project was set up in 1996, scientists have recruited many members of the public as volunteers. The most successful project in this category is SETI@Home (part of the Search for Extraterrestrial Intelligence), which was launched in 1999 (Korpela *et al.* 2001). As of December 2011, SETI@Home can count on almost 150 000 active volunteers, whose computers collectively contribute an approximate average of 550 TeraFLOPS, or 550 trillion floating point operations per second[4]. This rate of data processing is approximately equal to the world's third most powerful supercomputer.

The team behind SETI@Home developed a piece of middleware called the *Berkeley Open Infrastructure for Network Computing* (BOINC). BOINC provides a framework for scientists who wish to set up a CCP and also provides an interface which volunteers download in order to communicate with the project (Anderson 2004). One component of BOINC is a basic website through which a project can recruit and communicate with project volunteers. A particular feature of each BOINC project website are the forums: if a volunteer has difficulty with installing or running BOINC, or encounter difficulties with their work unit, they can seek advice on one of these forums, or they may use the forums to discuss general issues relating to the science underpinning a particular

[3] http://www.wxqa.com/ (accessed 9 December 2011).
[4] http://boincstats.com/stats/project_graph.php?pr=sah (accessed 9 December 2011).

project. The website also allows for the recording of statistics about the quantity of model data processed by project volunteers and their computers. A key statistic is known as *credit* (which measures the total quantity of data processed on an individual's computer). Credit scores are recorded and published on the projects' websites, with league tables ranking individual volunteers by credits obtained. These credits are very often a source of pride for project volunteers, who will sometimes report their number of credits in a signature placed at the end of forum posts. Furthermore, volunteers can group together and form teams, which also have credit scores (the credit score for a team for a particular project is simply an aggregate of the team members' individual credit scores) and ranking lists exist for teams as well. Very frequently, these teams have websites and forums of their own.

Another component of BOINC is the *BOINC interface*, which runs on the computer of each volunteer. It is through this that the volunteer communicates with the project: they use it to sign up for a particular project, download work units, be informed about the current progress of their work unit (for instance, what percentage of it has been completed, or whether it has crashed), upload completed work units to the project, and can monitor their credits. The interface also presents the volunteer with a list of all the other BOINC-based, and through this, the volunteer can register with other projects. It is possible for a volunteer to be active in a range of projects. As of December 2011, there are 11 projects using BOINC, in fields such as biochemistry, physics, mathematics, and climate science. Additionally, there are 61 projects using BOINC which are in the alpha or beta phase of development. Collectively, there are nearly 290 000 users of BOINC contributing a total of approximately 5 400 teraflops, a scale which massively outstrips that of any other middleware used in CCPs[5].

Nov *et al.* (2011) refers to such projects as *'low-granularity'* (p. 1), in the sense that the project requires a low investment from the volunteer in order to enable their participation. This contrasts with *'high-granularity'* projects (p. 2), where volunteers are asked to play a more active role, and it is to these that I shall now turn.

High-granularity projects: Volunteers play a more active role Some CCPs ask volunteers to play a more active role than simply downloading data onto their personal computers and having their computers process this data in the background. In these projects, volunteers typically visit a project's website where they are given tasks to perform. Such tasks might involve categorizing galaxies from telescope images, or looking for particles in microscope images, or finding conformations of folded proteins. The website www.distributedcomputing.info (accessed 9 December 2011) has identified approximately 30 such projects across a wide variety of scientific disciplines, including astronomy, bioinformatics and botany.

[5] www.boincstats.info (accessed 9 December 2011).

1.1.2 Improving the functioning of CCPs

There is recognition of substantial scope for improving the functioning of CCPs (Anderson & Fedak 2006), and there has been some work conducted on developing technologies to this end. In this subsection, I review this work, highlighting that not only is this seen as a technical challenge, but also as a social challenge by key individuals in the field of Citizen Cyberscience and they regard social studies of CCPs as making an important contribution in this respect. From this, my Research Questions are then motivated.

Technical challenges Work to address technical challenges for improving efficacy has been carried out by those already involved in CCPs, which suggests a strong commitment on their part to the continuation and strengthening of Citizen Cyberscience. Furthermore, such work has also been carried out by computer scientists who are not involved in any projects, thus reflecting a recognition by the broader academic community that Citizen Cyberscience is a significant field that requires support. For instance, in the case of the first category of projects outlined in the previous subsection, such work has included proposals that such projects should move from their current setup, where volunteers' computers communicate only with a project's central server (a *'client/server model'* (Pourebrahimi *et al.* 2005, p. 1)), to a more Peer-to-Peer-based network where volunteers' computers could communicate with each other (Costa *et al.* 2008). The idea behind this is to use spare bandwidth between volunteers' computers in order to reduce bottlenecking at the central server. Other proposals involve developing more lightweight middleware that might prove easier to use and does not require the same level of technical knowledge to implement as BOINC (for instance, Baldassari *et al*. 2006) and devising scheduling policies for the distribution of work units, in order to assign a project's work units to volunteers who have proved more reliable in the past (Kondo *et al*. 2007). These approaches all frame the task of improving the efficacy of Citizen Cyberscience as a technical challenge. However, this task can also be regarded as a social challenge, as I will now discuss.

Social challenges: Improving volunteer retention The success of CCPs requires the recruitment and retention of volunteers. As of December 2011, nearly five out of every six volunteers who have registered for BOINC projects no longer participate[6], suggesting possible scope for improvement. Increasing a project's volunteer retention rate should help in the pursuit both of increasing that project's scientific output and providing greater opportunities for members of the public to learn about the science underpinning the project.

[6] http://www.boinstats.info (accessed 9 December 2011).

To date, there have been three separate social studies of CCP, all of which address the question of how volunteer retention rates might be improved. These have been presented in Holohan & Garg (2005), Nov *et al*. (2010, 2011), and Raddick *et al*. (2010). The first in this list was something of a one-off paper, whilst the latter two report on long-term (and still ongoing) programs of research.

Questions may be raised about whether there is scholarly interest in the field of Citizen Cyberscience in social studies of such projects (and thus whether there would be any interest in the work described in this thesis), given how few have been carried out to date. However, it should be noted that two of these studies have been carried out in close collaboration with key individuals in the Citizen Cyberscience movement. For instance, Oded Nov has worked with David Anderson, who is the leader of, and most prominent spokesperson for, both the SETI@Home project and the team that developed BOINC. Indeed, Anderson is the one of three named authors (along with Ofer Arazy) on both of Nov's papers. Similarly, Raddick *et al*. (2010) is a collaboration with the leading members of another large-scale and well-established CCP, namely Galaxy Zoo. What these collaborative efforts suggest, then, is that not only are many leading and influential individuals within the field of Citizen Cyberscience (and across both categories of projects discussed in the previous subsection) interested in the question of how to improve volunteer retention rates, but that they recognise the potential contribution of social studies of CCPs to answering this question. Thus, I believe that the work presented in this thesis may prove of interest to those involved in Citizen Cyberscience.

Social studies of CCPs The three studies published to date have covered a range of CCPs, both high- and low-granularity, and have employed a range of qualitative and quantitative techniques. What is common, however, to all these studies is that they have focussed in particular on trying to understand what motivates volunteers to participate in CCPs.

The first published study, Holohan & Garg (2005), examined BOINC-based projects (low-granularity), applying both qualitative and quantitative methods to the studies of teams. By relating volunteer motivations to issues of organization, they presented the case that volunteers across the range of BOINC-based projects are primarily motivated by the credit system. In particular, they found that *'co-opetition'* (a mix of competition and cooperation) keeps volunteers engaged and, additionally, encourages them to contribute more to projects: the volunteers are competitive at an inter-team level in the sense of being keen to improve their team's league table positions in various projects, and cooperative within a team in the sense of providing encouragement and technical support on their team's online forums for their fellow team members.

Raddick *et al*. (2010) have focussed on a project of a very different nature, namely Galaxy Zoo (high granularity). Initially, they recruited 22 research participants at random from the whole corpus of volunteers and conducted interviews with these. A list of potential volunteer motivations (for instance, contributing to science, learning about astronomy, finding galaxy images aesthetically pleasing) was drawn from the transcripts of these interviews, which form the basis of an online survey administered on a large scale. Although the results of this survey have not yet been published, the interview data suggested that volunteers were motivated above all by the prospect of making a contribution to science.

The third study (Nov *et al*. 2010, 2011) also involves quantitative, large-scale surveys administered to volunteers in two projects, one being a low granularity project (SETI@Home) and the other a higher-granularity project (Stardust@Home). They presented each respondent with a list of possible motivations, and asked them to score each from 1 to 7 according to how important a factor they were in motivating the respondent. It was found that volunteers were particularly motivated by the prospect of contributing to the project's goals and the enjoyment they got out of performing the tasks required of them by the project. It was also found that there was a statistical correlation between this latter factor, and the task granularity, with a higher level of enjoyment recorded amongst volunteers involved in the high-granularity project.

Gaps in the literature and my Research Questions These studies have made an interesting start in increasing social scientific understandings of CCPs, covering a range of such projects and involving a number of research methods. Nevertheless, it is inevitable that there are significant gaps in this literature, given that there have only been three social studies of CCPs to date. I will highlight two in particular, and it is to address these gaps that the three questions which have guided my research have been formulated.

The first is that none of the three studies have problematized either the role of a project's scientists and software engineers or the project that they have built (e.g. the interface through which the volunteers can communicate with the scientists, the content of the project website, the methods by which volunteer contributions are acknowledged etc.). A project's scientists and software engineers are social actors, possessing and pursuing interests. Furthermore, they can exercise a great deal of discretion when deciding how to design and develop their projects in order better pursue their interests. As Winner (1980) argues, items of technology can have politics embedded in them by their designers. For example, methods of acknowledging volunteer contributions may be employed by a project's team members in order to attempt to control and direct the project participants. The way in which they design such methods communicate to the project volunteers the way in which they want the volunteers to behave: in the case of SETI@Home, for example, credits are awarded only once a work unit is completed, thereby attempting to provide an incentive for volunteers not to abandon their work unit before completion (Anderson 2003). The project team also communicate with volunteers through the content they choose to add to (or withhold from) the project's website. How they behave, and the decisions they make, can thus have a potentially powerful impact upon volunteers experiences (and thus motivation to continue participating) in a project.

Based on the above discussion, the first two Research Questions are motivated. These are:

Research Question 1: How do scientists and software engineers who are about to embark upon CCPs intend to establish, structure, and maintain relationships with members of the public?

Research Question 2: How do those with some experience of setting-up and administering CCPs seek to establish, structure, and maintain relationships with project volunteers?

The second gap in the literature regarding social studies of CCPs that I shall highlight here is that the studies conducted to date have tended to regard volunteers as a single group, who may not be homogeneous but whose characteristics are treated as varying along a continuum, and thus can be captured in a single set of descriptive statistics (means, standard deviations etc.). However, as Michael (1992) argues, laypeople are not simply a single group, but are heterogeneous, consisting of different groups with a multiplicity of (sometimes conflicting) perspectives, backgrounds, and interests. It is reasonable to assume that the lay volunteers of a CCP, too, are heterogeneous and may seek to position themselves with respect to each other or come into conflict. The dynamics amongst the various volunteers may therefore have an impact upon volunteer experiences (and thus motivation to participate) in a project. Volunteers may construct relationships with each other in the online forums, for instance through the content of posts they make (what they include in the posts and what they withhold, the tone they adopt in their posts, whose messages they choose to respond to and whose they ignore and so on). For instance, a volunteer who seeks to establish and maintain a particular position in a social hierarchy on the project forums might attempt to do so by posting messages in which they seek to establish which topics they have authority to respond to and how others should relate to them.

Thus, my third Research Question is motivated:

Research Question 3: How do project volunteers seek to establish, structure and maintain relationships with other project volunteers?

Here, then, the three Research Questions that guided the research presented in this thesis have been presented. As I will discuss in Chapter 3, a bottom-up, qualitative ethnographic approach to studying CCPs was a sound choice in order to answer these Questions.

In this Section, the Research Questions have been justified by arguing that social studies of CCPs are already of great interest to many leading figures within Citizen Cyberscience, that the existing studies have left gaps in the literature, and that the Research Questions should help to address these gaps. However, as I will argue in the next Section, the answers to these Research Questions should also help to address gaps in the Science and Technology Studies literature and prove of interest to scholars in this field.

1.2 Motivations for this thesis II: Contributions to Science and Technology Studies

In addition to providing the basis for recommendations for improving the running of CCPs, it is also anticipated that the work presented in this thesis will make a substantive contribution to the field of Science and Technology Studies (STS), in particular, that its conclusions may prove of interest to the large body of social scientists who study interactions between members of the public and technoscience and amongst members of the public around scientific issues (for instance, see Chilvers 2008).

The answers to RQ1 and RQ2 should help to advance the understanding of interactions between members of the public and professionals/experts in the domains of science and technology (I shall henceforth refer to this latter group as *technoscientists*). To summarize rather broadly certain arguments that will be presented in the Literature Review (Chapter 2), the existing studies of such interactions can be largely sorted into three categories: it is my contention that the research in this thesis does not fit into any of these categories, as it describes situations where technoscientists, having been thoroughly embedded in a scientific culture, are engaging with members of the public around scientific issues, giving rise to emergent relations between technoscientists and members of the public.

To see this, I shall outline the three main categories. The first concerns situations where technoscientists doggedly attempt to construct and maintain boundaries between themselves and members of the public in order to exclude these members of the public from making substantive contributions to technoscientific decision-making processes. Many studies which fall into this category are analyses of public inquiries, or public engagement mechanisms. The second category involve situations where certain lay members of the public or lay groups have been fully-integrated into, and make substantive impacts upon, at least some stages of technoscientific decision-making processes. Here, it appears that some technoscientists have allowed the lay-science boundary to be breached, but in reality, the members of the public allowed to breach this boundary have adopted rationalist discourses of science, and hence technoscientists have been able to continue to construct a boundary, conferring legitimacy to contribute on those within the boundary (i.e. those who deploy a scientific discourse) and denying this legitimacy to those outside the boundary. This apparently occurs when certain lay people feel a desperate need to impact upon these processes (for instance, Epstein 1996). The final broad category are situations where the power relations between the technoscientists and their representatives, and the lay public are very much reversed, and it is very much in the interests for those technoscientists interacting with the public to adopt the discourse of members of the public.

There therefore appears to be a gap in the literature, in that there are few accounts of emergent relations between members of the public and technoscientists. Nevertheless, it seems that such accounts could be of value to those seeking to design and assess public engagement mechanisms which seek to allow laypeople to have a substantive impact upon technoscientific decision-making processes. Case studies of these mechanisms have suggested that, whilst laypeople's voices have been heard in such processes, it is often questionable whether they have had any substantive impact (e.g. Kaminstein 1996). As a result, there has been an increasing move towards ensuring that laypeople's viewpoints are incorporated in decision- and policy-making (Nowotny 1981), including calls to make this legally enforceable (for instance, see Rowe & Frewer 2004), and, to this end, evaluations of these public engagement mechanisms have increasingly involved attempts to measure the extent to which they are impacted by lay perspectives (e.g. Einsiedale *et al.* 2001). Understanding how technoscientists with little previous experience of engagement with the public around scientific issues might evaluate and incorporate these perspectives will therefore be of increasing importance as public engagement mechanisms are subjected to greater scrutiny to ensure this occurs. There is a gap in the literature regarding this; it is hoped that my answers to RQ1 and RQ2 will begin to fill it.

I also hope that the results of my analyses which are directed towards answering Research Question 3 will prove of use to those seeking to design public engagement mechanisms which will result in distributive justice for those affected by the outcomes of various technoscientific decision-making processes, because I aim to provide insights into how certain lay people involved in CCPs attempt to construct and manage social relations with other volunteers, casting them as possessing or lacking the legitimacy to contribute to various aspects of technoscientific decision-making processes within the project. As is noted in the Literature Review, however, there has been very little attention indeed paid to this and therefore conclusions from my case studies will, I hope, begin to fill this gap.

Attaining distributive justice is not merely a case of ensuring that technoscientists allow the views of affected laypeople to have a genuinely substantive impact upon the outcomes of these processes. It must also be recognized that the interests of various lay groups and individuals are often likely to conflict during the course of such processes, and as a result inequalities of power between these lay groups may lead to socially unjust outcomes where one lay group is able to marginalize, force out, or misrepresent another lay group (for instance, see Epstein 1996). There is good reason to believe that various lay groups will seek to advance their own interests during the course of technoscientific decision-making processes, possibly at the expense of other lay groups: for example, a lay group may attempt to construct themselves as more scientifically capable than a rival group in order to recruit allies from amongst the participating technoscientists. Understanding how various interested lay groups may seek to do so, is therefore critically important for designing public engagement mechanisms which are socially just.

1.3 Case studies

In order to answer my research questions, qualitative ethnographic case studies of two CCPs have been conducted. Following Knorr-Cetina (1999), I felt that conducting two in-depth case studies of projects in separate scientific disciplines would allow for the highlighting of differences as well as the drawing out of similarities in order to understand how the projects proceeded according to the different (institutional, political, historical) contexts in which they are situated. In Chapter 3, I describe my research methodology (including justification for my underlying approaches) in greater depth.

These projects are climateprediction.net (in the field of climate science) and Galaxy Zoo (in the field of astronomy). These two projects were selected because it was felt that they complement each other well, in that they are sufficiently different that they would allow for comparisons and contrasts to be drawn between them: climateprediction.net is a BOINC-based project and as such is a high-granularity project, whilst Galaxy Zoo invites volunteers to play a more active role by presenting them with pictures of galaxies and asking them to categories them and thus is a high-granularity project. Furthermore, both are well-established projects (climateprediction.net was launched in 2002, whilst Galaxy Zoo commenced in 2007), each with a large number of active volunteers (approximately 30 000 in the case of climateprediction.net and 150 000 in the case of Galaxy Zoo). This meant that each project promised to be a very rich site of interactions which I could study. Furthermore, the maturity of the projects allow for greater reflection on the part of stakeholders about their involvement in the project. Both projects have developed to the point where lessons learned are being reincorporated into future development, and thus promised to provide a strong empirical basis for the formulation of recommendations for the running of CCPs.

climateprediction.net The climateprediction.net project was set up, and is run by, a team of atmospheric physicists and computer software specialists at the University of Oxford. It has three main aims:

- To produce scientific results, in particular to contribute towards improved understandings of possible future climate change that might, for instance, be useful in policy-making processes;
- To demonstrate the suitability of a Citizen Cyberscience approach to conducting scientific research involving multiple runs of simulations; and
- To educate the public about climate science and increase public support for the field, and for policies to mitigate potential future climate change.

The project began in 1999, and in 2002, the decision was taken to transform it into a CCP. Following alpha and beta testing in late 2002 and spring 2003, climateprediction.net was launched to the public in September 2003. In August 2004, the project moved to using BOINC, and continues to use this middleware. As of December 2011, there are approximately 26 000 active users, who download and run models simulating and forecasting global climate change. This project, and the particular historical, scientific, and social contexts in which it is embedded, will be discussed in more detail in Chapter 6.

Galaxy Zoo The Galaxy Zoo project is run by a consortium of astronomers and software developers based in a range of scientific institutions. Its two current main aims are to educate members of the public about astronomy and to produce scientific results, in this case classifying galaxies according to shape. It has been found that the human eye far outstrips any automated computational approach in terms of classifying galaxies, yet the sheer number of galaxies means that it is not feasible to rely on professional scientists alone to perform this task. This therefore seems a suitable basis for a CCP.

Galaxy Zoo is currently in its third phase. Its first phase was launched in July 2007, and recruited over 100 000 members of the public to classify approximately one million images of galaxies produced by the Sloan Digital Sky Survey. The second phase of Galaxy Zoo, Galaxy Zoo 2, was launched in early 2009, and asked volunteers a more detailed series of questions about a quarter of the original images. The third phase, launched in mid-2010, asks volunteers to classify images taken from the Hubble Space Telescope. This project, and the contexts in which it is embedded, will be described in greater detail in Chapter 4.

1.4 Looking ahead to the rest of this thesis

So far, this introductory Chapter has motivated this thesis in terms of its potential contribution to improving the running of CCPs, as well as suggesting how it might help to advance understandings of relations between scientists and members of the public, and amongst members of the public, around issues of technoscience. There are six Chapters that follow this Chapter. In ***Chapter 2***, I review the social science literature relating to interactions between scientists and members of the public, and between members of the public over issues of technoscience. This review identifies gaps in the existing literature, and explains how the work of this thesis will make a contribution to filling these gaps. I also discuss how this existing literature is of only limited applicability to the development of recommendations for volunteer engagement in CCPs, and thus the work presented in this thesis should therefore be of great use in developing such recommendations.

Chapter 3 describes the methodology I employed to set about answering my Research Questions. I justify my choice of a qualitative, ethnographic approach to studying CCPs, and I discuss some of the issues that might arise when applying such an approach to the study of communities that are mediated primarily by the Internet. I then explain how I went about collecting my data, including how I took into account ethical issues involved, before discussing how I analysed this data.

I present my findings from the Galaxy Zoo case study in ***Chapters 4 & 5,*** which relate to the relationships between scientists and volunteers in the Galaxy Zoo project, i.e. they relate to RQs 1 and 2. I discuss the ethos that has emerged in the project regarding how the project team members should treat the volunteers, and the conditions under which objects and knowledge are exchanged amongst a project's scientists and volunteers. This argument is split across Chapters 4 and 5 as follows: in Chapter 4, I introduce the project and consider the broader contexts into which it has emerged, before considering the interests of the team involved in the setting-up and running of Galaxy Zoo in these contexts. In Chapter 5, I will then describe the key elements of the ethos regarding how Galaxy Zoo scientists should treat volunteers. By drawing on the conclusions of Chapter 4, I discuss how this ethos has emerged as a result of choices which the core team believed would better enable themselves to pursue their interests. Finally, I argue how the features and policies considered – and the discourse surrounding them – constitute a particular form of knowledge economy that governs the conditions of exchange amongst volunteers and project scientists.

I then turn to my case study of climateprediction.net in ***Chapters 6 & 7.*** The aims are two-fold: the first is to clarify my analysis in Chapters 4 and 5 (and thus help to answer RQs 1 and 2) and to see how some groups of volunteers attempt to position themselves relative to each over (RQ 3). I examine how the credit and other statistics systems climateprediction.net have played a critical role in the project as experienced by both volunteers and project team members. Chapter 6 focusses on these team members. First, I situate the project in its broader contexts, in order to understand how the project emerged and how its goals are oriented towards these contexts. I then explain the major features of the project, and how the project has developed over time. This sets the scene for charting the team members' changing perceptions of the volunteers, in particular how, as the project unfolded, these team members have developed a strong belief that the statistics systems is the critical factor in maintaining volunteer interest in the project. In Chapter 7, I consider the significance of the statistics systems for the volunteers. During the course of my study of climateprediction.net, a number of different groups of volunteers have been identified, based on the extent and nature of their participation in the VCPs. In Chapter 7, I compare and contrast two of these classes. In particular, I argue that the statistics systems play very important, but very different, roles in each group's experiences of climateprediction.net. This complements the conclusions of Chapter 6 by demonstrating that the social reality of these systems vary not only across temporal contexts but also across spatial contexts.

Finally, in ***Chapter 8***, I draw together the empirical work presented in Chapters 4 to 7. I reflect upon the implications of my findings in relation to the literature reviewed in Chapter 2, and consider how CCPs might be designed in future and recommendations for practice for people who are involved in running, or interested in setting up, such a project. I finish by discussing how the work presented in this thesis might be extended.

Chapter 2: A review of the literature on public engagement with science

In this Chapter, a review is presented of the social science literature which deals with how scientists and non-scientists interact. In particular, the extent of this literature's potential relevance to answering Research Questions 1 and 2 (RQ1 and RQ2) will be considered: similarities between the situations in the case studies in this literature and the challenges facing scientists and software engineers in CCPs will be highlighted, but differences which might limit the relevance of insights from these case studies will be drawn out. Also detailed will be attempts to find literature relevant to Research Question 3 (RQ3), namely how various non-scientists interact with each other around issues relating to technoscience: it will be noted that such literature is very scarce indeed, and an attempt to account for why this is the case will be presented.

The studies of interactions and relationships between scientists and non-scientists are, broadly, grouped into three categories in this review, depending on the nature of these relationships that they present. One of the major conclusions of this review is that the nature that these relationships take seems to be strongly related to the interests of the parties involved. These parties appear to employ a range of discursive moves which define how these interactions take place in ways that seem to allow them to pursue their interests: these moves cast themselves and others in particular roles through their uses of language, whose utterances they acknowledge or ignore, to whom they address their utterances and the form these utterances take, and so on. It is these discursive moves that enable the people making them to pursue their interests: for instance, they may have the effect of silencing or marginalizing those who might be a threat to their interests, or to enroll others as allies.

The first category of studies presented involve instances where technoscientists construct relations with laypeople whereby they seek to define themselves as sole possessors of legitimate knowledge, generated by rational, scientific processes, and to cast any dissent or resistance on the part of laypeople as stemming from ignorance or irrationality (this has been given the term *'deficit model'* (Miller 2001, p. 116)). It is argued that this is a result of technoscientists pursuing a strategy to construct and maintain a boundary between themselves and non-scientists, in order to preserve the position of science as the preeminent knowledge-generating institution in society. It is also argued that such an approach can result in the alienation of non-scientists from scientific institutions, with lay people feeling ignored and patronized, because it actually miscasts both the way in which the knowledge of the technoscientists is generated and also refuses to acknowledge the complexity and validity of the knowledge possessed by laypeople, whose dissent or rejection of technoscientists' arguments and advice is not simply on the basis of ignorance or irrationality. In the situations presented in this first category, the technoscientists believe they have little to lose from alienating the lay people with whom they are interacting and a great deal to gain from drawing and reinforcing boundaries between themselves and others, leading to situations where there is little or no engagement of laypeople by scientists, or vice versa.

The second and third categories contain cases where some engagement has taken place. Those cases in the second category are characterized by situations where some laypeople have engaged technoscientists by adopting scientific discourses, in the sense that they use the same terminology and style of argument as the technoscientists: in these situations, the technoscientists still seek to reinforce the boundary between science and the rest of society, but allow themselves to become engaged and enrolled by particular laypeople when these laypeople are willing to adopt the discourses which maintain this boundary. As shall be argued, in these situations, technoscientists will seek to defend and advance the interests of science as an institution, similar to the first category, but certain lay people have a very strong interest indeed in engaging these technoscientists in order to influence the way they act. For instance, one case study details how desperately ill HIV/AIDS patients and their advocates in the 1980s sought to change the way in which clinical trials of new treatments were conducted in order that they might gain access to the trials (and hence the new drugs), and some achieved this through learning the knowledge and discourses of medical science so that they might engage medical researchers and policy-makers on their own terms (Epstein 1996). In the third category, the situation is reversed: in the cases here, the technoscientists and their representatives appear to acknowledge the validity and complexity of the knowledge and views of laypeople, and shape their discourses to accommodate these in an attempt to engage and enroll members of the public. Many of these case studies are of how science is represented in media outlets or in museum exhibits, where those presenting accounts of science to the public have interests (such as the need to sell newspapers or attract visitors to a museum) which would be jeopardized by alienating the public, and which may therefore prove more compelling than attempting to construct and maintain a boundary between science and the rest of society.

The analysis presented in all of these case studies draws heavily upon insights gleaned from the field of *Science and Technology Studies* (STS) regarding the social processes by which knowledge-claims presented by scientists are generated, contested, and eventually either accepted as true (and thereby added to the body of widely-accepted scientific knowledge) or deemed false. Before reviewing the literature regarding interactions between scientists and non-scientists, an overview of some of the major theoretical components of STS will be presented, particularly focusing on those relevant to the subsequent literature review.

2.1 Science and Technology Studies (STS)

Practitioners within the field of STS took a critical perspective on the so-called *'canonical view'* of such processes that was dominant until the mid-1970s (Bucchi 1998, p. 4), a view underpinned by the notion that there are certain essential features of scientific practice which set it apart from other methods and practices by which knowledge claims about the world are made (such as astrology), and which ensure that the processes by which science generates knowledge are objective and rational.

Instead of this canonical view, practitioners of STS have set out to show how processes by which scientific knowledge is generated and certified as true are contingent upon the local material and social contexts where these processes take place, and that these processes are governed by the pursuit of interests (for instance, social status or economic interest) of those actors involved in these processes. In this Section, I first set out the essentialist canonical account of the scientific method before turning attention to some of the work conducted within STS, in particular highlighting insights from two subfields of STS. One is the *Sociology of Scientific Knowledge* (SSK), which sought to show how the content of scientific knowledge is determined by social factors as individual actors seek to pursue their own interests. The other is *Laboratory Studies*, which attempted to demonstrate the highly contingent and contextual nature of the processes by which scientific knowledge is generated. Finally, the concept of *epistemic cultures* is presented, which will prove particularly relevant in understanding the theory underpinning many of the studies of interactions between scientists and non-scientists presented in Section 2.2 below.

2.1.1 Canonical account of science

In the early 1960s, the canonical view regarding natural science was that it was about the progressive accumulation of objective knowledge, consisting of a collection of theories which were grounded in real-world observations. According to this view, scientists systematically collect data about natural phenomena based on experiments, either in a laboratory or in the field. From careful analysis of this data, the scientists then derive theories which they claim describe the world. These claims are then open to scrutiny by other scientists: those advancing knowledge claims are expected publish not only these claims but descriptions of their method so that their experiments may be replicated, and verified, by others (Collins 1985). Claims which are then seen to have proven themselves in the face of scrutiny are then elevated to the status of fact.

This canonical view was rooted in the objective/subjective dualism: science was perceived as *'establishing its knowledge-claims by objective and absolute methods'* (Dolby 1971, p. 3), whereas scientific error (for instance, theories which were proposed but were later generally regarded as false) or dissent to parts of scientific knowledge were seen as the product of irrational biases, prejudices, or personal interests held by those who advanced such theories or dissent.

Many attempts were made to find the essence of science, to draw out a set of criteria against which knowledge-building practices could be judged to determine whether or not they were scientific. It was widely held that such a set of demarcation criteria could be found; the only challenge was to establish what they were (for instance, see Popper (1959) for a set of epistemological demarcation criteria; see Merton (1949) for a set of demarcation criteria derived from sociological analysis).

The implications of this approach for the social studies of science were threefold. Firstly, science was regarded as a unified endeavour. Although scientists may be divided according to discipline or institution, demarcation criteria were believed to apply universally and hence scientific knowledge was being generated, scrutinized and certified according to the same standards irrespective of discipline or institution. The second, related, implication was that it meant that such studies were largely restricted to the study of science on the level of society-at-large, the study of how whole disciplines, scientific institutions and university departments emerged, existed, and sometimes disappeared. The impact of society upon natural science was seen as being restricted to determining which disciplines received the greatest public support, and hence public funding. The content of the knowledge generated by a particular discipline was not seen as being fundamentally affected by the amount of funding it received or the rate at which it grew; rather, such issues would affect simply the rate at which the discipline accumulated its knowledge (Cole & Cole 1973).

In the early 1970s, however, some sociologists began to question these approaches to the study of science (Barnes & Dolby 1970; Dolby 1971; King 1971). For instance, Dolby (1971) argued that proposed demarcation criteria, far from describing the essence of what natural science was, were instead *'excessively unrealistic idealizations'* (p. 7). These sociologists were heavily influenced by historian of science Thomas Kuhn's work, *The Structure of Scientific Revolutions* (Kuhn 1962), and key ideas presented therein formed the basis of a new approach to social studies of science, the *Sociology of Scientific Knowledge* (SSK).

2.1.2 Thomas Kuhn and SSK

Kuhn rejected the canonical view of scientific practice, namely that the body of scientific knowledge grows incrementally and involves approaching new data with an unbiased mind, instead arguing that scientists bring a pre-existing body of theory to which they attempt to fit this data (Kuhn 1962). Kuhn labels such a body of theory a *'paradigm'* (*Ibid.*, p. 10. By surveying history, Kuhn argued that there have been long periods of relative stability in the scientific endeavour (known as periods of *'normal science'* (p. 10)), during which the same scientific paradigm persisted over a long period of time, punctuated by sudden shifts in paradigms, where the existing paradigm was abandoned as a result of an increasingly intolerable body of anomalies (where data could not be fitted to the theory of the dominant paradigm), with a consensus forming instead around a different body of theories. Such shifts refer to the *'scientific revolutions'* in the title of Kuhn's book.

An example given by Kuhn of shifting paradigms within a field is in optics, and the way in which light was characterized within each paradigm: in the eighteenth century, the paradigm was that light consists of material corpuscles (i.e. very small physical objects); this gave way to the characterization of light as a wave; in turn, following the work of quantum physicists, the paradigm emerged that light consists of entities called photons, which exhibit some characteristics of waves and some of particles (*Ibid.*).

The details of Kuhn's descriptions of how these revolutions proceed will not be dwelt on here: the aim instead is to highlight the concept which Kuhn calls *'incommensurability'* (*Ibid.*, p. 103) between different scientific paradigms because it is through the exploration of the implications of this concept that the major theoretical foundations of SSK were laid down. According to Kuhn, any two paradigms are incommensurable. A paradigm change involves a thoroughgoing re-evaluation through the lens of the new dominant body of theories not only of the body of scientific knowledge but also *'back upon the science that produced them'* (*Ibid.*, p. 149), leading to marked changes in the standards by which solutions to problems are assessed, the methods used in obtaining these solutions, and the field of problems whose solutions are considered important. Although new paradigms often employ the same vocabulary and apparatus as old ones, they *'seldom employ these borrowed elements in quite the traditional way'*; instead *'old terms, concepts and experiments fall in to new relationships one with the other'*, leading to misunderstandings and only partial comprehension between scientists working within different paradigms (*Ibid.*, p. 149).

Kuhn highlights some implications of incommensurability, and one in particular played an important foundational role in SSK, namely that *'theory must be chosen for reasons that are ultimately personal and subjective'*, because there is no possibility that choices between competing theories from different paradigms can be made on the basis of *'good reasons'* alone (*Ibid.*, p. 199): any attempt by a scientist working within one paradigm to make a compelling case based entirely on appeals to objective standards of truth to a about the superiority of her theoretical paradigm to a scientist working within a different paradigm founded on a different body of theories is bound to fail.

Early proponents of SSK developed this insight further, claiming that any judgment made by a scientist about knowledge claims and theories advanced by others was ultimately determined by personal and subjective reasons. Any two scientists - even those working at the same time within the same scientific institutions, or even at the same bench in the same laboratory - work within different paradigms and that the paradigm within which an individual was working was continuously liable to change. This is because the precise meanings of concepts (theories, classification systems, labels etc.) used by scientists to interpret a novel situation are never completely determined in advance of the situation and independent of those who employ them; instead, their application in a particular instance is *'developed step by step, in processes involving successions of on-the-spot judgments'*, and hence their meaning varies from individual to individual and from situation to situation (Barnes 1982, p. 30).

The implication of this is that, it was argued, the success of a particular scientific theory in becoming accepted as scientific truth can be accounted for in sociological terms by considering the interests of various actors involved in scientific controversies and debates, rather than simply explaining this theory's success as a result of it accurately representing nature (Barnes *et al.* 1996; Collins 1985). Shapin (1979) expands on the notion of how social factors, and the social interests of individuals, drives theory choice, explaining that *'people, universally, have an interest in social management and control. They try to maintain institutions that serve them well; or they actively try to criticize and discredit institutions which work against their interest'* and, critically, that these goals *'have a bearing upon the processes via which the knowledge is constructed, judged, and institutionalized'* (Shapin 1979, p. 46, after Douglas (1966)). In other words, when presented with competing theories and knowledge claims, an individual - even a scientist - will choose according to which theory or knowledge claim best suits the pursuit of their interests.

To illustrate this point, a number of case studies of specific scientific controversies were carried out which linked the theory choice made by actors to their social interests. For instance, in a study of the debate in early 19th century Edinburgh regarding phrenology (the idea that particular psychological features of individuals can be related to the shape of particular parts of their brain), Shapin (1979) argues that members of the bourgeois and petty-bourgeois (the elites who had traditionally dominated Edinburgh social life and whose position in the social order was under threat from the industrial classes) supported phrenology because it lent itself to supporting the argument that God had preordained the characteristics of individuals, justifying the perpetuation of a natural social order, and hence helping to maintain their elevated social position.

The notion that theory choice is driven by social interests was extended by studies carried out of the processes occurring at the site where knowledge is first generated, namely the laboratory.

2.1.3 Laboratory Studies

These studies were based upon ethnographic research, which involved close observation of scientists and technicians in a single laboratory, studying their work practices, the equipment they constructed and used, how physical objects under study were handled, how events in the laboratory were discussed and interpreted, and the various products of processes within the laboratory (notable examples include Knorr-Cetina 1981; Latour & Woolgar 1979; Lynch 1985; Traweek 1988).

There are two key insights from laboratory studies which will be highlighted here. The first is that knowledge production (such as the interpretations of experimental observations) by scientists in the laboratory is oriented towards the acquisition of credibility for the scientists in the social worlds outside the laboratory. In *Laboratory Life*, Latour & Woolgar argue that the creation of *'order out of disorder'* (Latour & Woolgar 1979, p. 235), to derive a plausible scientific account from the seemingly chaotic and complex natural processes being studied, was driven by the ultimate goal of producing publications (journal articles, conference papers) which would prove robust in the face of opposition from scientists working in other laboratories (see also Latour 1987, ch. 2). Every action or occurrence in the laboratory seemed to be recorded in some way (*'inscriptions'* such as: laboratory notebooks, labels written on test tubes, images produced by laboratory equipment and annotated by scientists, graphs and other visual representations of data, drafts of scientific papers), and the form that these inscriptions took were determined completely by their anticipated usefulness in formulating articles intended for publication (Latour & Woolgar 1979, p. 48).

The second key insight is that the content of knowledge produced within a laboratory is also influenced by the immediate context (the social order within the laboratory) as scientists wish to gain credibility and status within the laboratory, and is also contingent upon chance events within the individual laboratory. Such events might include a scientist's unexpected encounter with another, who might suggest a new interpretation for data which sets the scientist off down a new path of interpretations which would otherwise have remained unexplored (Lynch 1985). Even the physical layout of the laboratory can impact upon the frequency of such encounters: for instance, two individuals will interact with a different frequency if they are on adjacent benches than if they are located in different rooms (Latour & Woolgar 1979). Another event might be a particular piece of equipment becoming unexpectedly available to a scientist, which in turn might lead to a whole succession of experimental work being carried out which would not otherwise have been undertaken (Knorr-Cetina 1981).

Together, insights from SSK and Laboratory Studies suggest that scientific knowledge can be seen as contingent, in the sense that the processes by which it is generated, contested, and certified as true knowledge, are driven by the social orders that exist within the immediate local settings where they occur (such as the laboratory) and in society-at-large. It can also be said that the content of this knowledge is highly contingent, both upon these social contexts, but also upon the occurrence of chance events.

Before introducing the concept of epistemic cultures, two other developments within the field of STS will be briefly highlighted and will be useful in understanding the meaning of this concept. The first are attempts to reinsert non-social factors in accounts of how scientific knowledge is generated. One example is the approach known as *Actor-Network Theory*, which involved the dissolution of boundaries between humans and non-humans, with the assignation of interests to non-humans and the assumption that these non-humans would act in ways to safeguard or advance these interests, in turn impacting the production and contestation of knowledge claims (for instance, see Callon 1986a, 1986b; Latour 1987, 1988, 2007). Another attempt to reintroduce non-social factors has been made by Pickering (1995, 2006), who has argued that scientists are motivated and constrained by a number of factors (social, material, logical).

The other development within STS is the recognition of the fundamental role played by trust in processes of generation and certification of scientific knowledge, and, furthermore that knowing who to trust, and in which situations, is an act of knowledge-building in itself, and hence, the social interests of individuals will play an important role in determining whose accounts of the world these individuals wish to trust (Shapin 1994). It was noted that the work of a scientist, or team of scientists, does not take place in a vacuum. Instead, when performing their work, these scientists will take as true theories generated by other scientists, in other laboratories and, frequently, in other disciplines, and will use them to interpret what they observe in their own experimental work (Hardwig 1985, 1991). Trust also underpins collaborative research work in science, the prevalence and scale of which has been increasing rapidly in recent decades (Galison & Hevly 1992). For instance, in some fields, it is common for the number of collaborators in a single project to extend into the hundreds, and thus a single scientist in such a project cannot verify the results presented by each of their collaborators (e.g. Knorr-Cetina 1999). As a result, *'scientific propositions often must be accepted on the basis of evidence that only others have'* (Hardwig 1991, p. 706).

2.1.4 Epistemic cultures

Here, the notion of culture will be dealt with first, before specifying the meaning of the term *epistemic cultures* and how this term relates to the insights from STS highlighted above. Geertz (1973) defined a culture as *'a historically transmitted pattern of meanings...a system of inherited conceptions expressed in symbolic form by* *means of which men communicate, perpetuate, and develop their knowledge about and attitudes towards life'* (p. 89). Culture results from the employment of strategies by people to try and make sense of, and bring order to, the world around them, and the subsequent learning of these strategies by others within a particular community to the extent that they become shared practices (customs).

Customs arise because people need to cooperate in order to make sense of the world around them so they can pursue their ends, and hence need to behave in ways that are meaningful to others. Culture, then, *'means a thick growth of variegated patterns piling up on top of one another...involving multiple instrumental, linguistic, theoretical, organizational, and many other frameworks'* (Knorr-Cetina 1999, p. 10).

Different cultures are incommensurable. The development of a culture is contingent on the local contexts in which it arises. In this sense, two individuals from different cultures can only achieve a partial shared understanding: even if they share a vocabulary, the meanings they assign to particular words, and the way in which they relate words to each other, will be shaped by the differing cultural contexts in which they have become familiar with these words (Strathern 1992).

Knorr-Cetina defines *'epistemic cultures'* as *'cultures that create and warrant knowledge'*, namely *'those amalgams of arrangements and mechanisms - bonded through affinity, necessity, and historical coincidence - which...make up how we know what we know'* (Knorr-Cetina 1999, p. 1). An epistemic culture contains customs regarding: procedures to generate knowledge; the form of the evidence that should be presented in support of particular types of knowledge claims; linguistic norms regarding the forms that statements of knowledge should take, whose accounts of knowledge should be trusted (see also Shapin 1994) and whose accounts should be dismissed; how producers of knowledge should be rewarded; how those involved in knowledge production should relate to each other from a hierarchical point of view; and who should be accorded power to determine whether a particular knowledge claim should be accepted as true.

It is unsurprising to find that distinct epistemic cultures have arisen in different scientific subdisciplines (Knorr-Cetina 1999; see also Galison & Stump 1996). For instance, the expectations that society has of scientists will differ according to discipline (for example, the expectation that the knowledge generated by nuclear physics will impact upon the development of nuclear power; the hope that the work of medical science will lead to cures for diseases and so on), and so, in turn, the standards by which the knowledge produced is judged and rewarded will differ according to discipline, which will impact upon scientists' practices as they orient their practices towards maximizing their rewards (Forman 1971, 1987).

Additionally, the differences between disciplines in terms of the physical phenomena they study also drives the development of distinct cultures, giving rise to the use of different instruments and practices to manipulate and record data, differing patterns of collaboration and differences to the extent that scientists are required to trust in others' knowledge claims. Hence, both the differing local conditions of knowledge production within, and the variation in expectations that society have of, different disciplines give rise to distinct epistemic cultures. These cultures can be said to be incommensurable. Knowledge claims in one discipline will be generated according to different methods, modes of reasoning, and patterns of trust than those in other disciplines, and will be evaluated and accepted or rejected in different ways.

A number of theoretical insights from the field of STS have been presented in this Section that together suggest that the generation of scientific knowledge is not a completely objective and unified endeavour. Rather, the processes involved are of a contingent and contextual nature, in which the pursuit of interests by actors, chance events, and the epistemic cultures (which can vary between scientific disciplines and institutions) in which actors are embedded all play a significant role. According to STS, there are no essential features of science which demarcate it from non-science and ensure that the processes by which scientific knowledge is generated and knowledge controversies are resolved are free from social commitments and purely objective; instead, scientists must engage in *'boundary work'*, namely strategic (such as rhetorical or discursive) moves in order to maintain the cultural boundary by which their institutions are separated from rival institutions making knowledge claims, and between the institution of science and the rest of society (Gieryn 1983, p. 781. See also Beatty 2006; Burri 2008; Gieryn 1999). It is with this in mind that consideration turns to the review of the literature regarding interactions between scientists and members of the public because this literature draws heavily upon the theoretical orientations provided by STS.

2.2 Interactions between scientists and the public

The purpose of this Section is to provide a review of the literature that deals with interactions between technoscientists and members of the public, particularly focusing on how those involved seek to structure and maintain relationships with each other, and how these processes relate to the pursuit of their various interests. This Section presents case studies regarding these relationships, grouping the literature into three categories according to the nature of the relationships established, and is particularly relevant to RQ1 and RQ2 (literature relevant to RQ3 will be considered in Section 2.3).

2.2.1 Lay knowledge and the *'deficit model'*[7]

The first group of studies considered comprises case studies where relationships between laypeople and technoscientists have broken down, as a result of boundary work carried out by technoscientists whereby they alienate laypeople and appear to have little interest in repairing the relationships. The analysts who present these case studies seem to suggest that the technoscientists have engaged in this to protect science's authority as the pre-eminent knowledge-generating institution within society, and that they have done so by seeking to construct their relationships with laypeople in this way in order to protect science's authority through making strategic moves to create a cultural boundary between themselves and laypeople, and that alienating these laypeople is a price worth paying for this.

Two aspects of these case studies will be explained below. The first is how these studies detail the moves they make in constructing their interactions with laypeople, including the assumptions embedded in these moves regarding laypeople's ability to handle scientific knowledge, and may therefore be directly relevant to RQ1 and RQ2. The second aspect is how such constructions alienate laypeople, particularly focusing on how the assumptions embedded in these constructions regarding how laypeople handle scientific knowledge do not capture reality. This will prove critically important in understanding literature in the later Sections of this review when case studies of different types of relationships between technoscientists and laypeople are presented. Before presenting the case studies in this Section, the historical context in which these studies emerged shall be outlined: those conducting these case studies shared a common agenda, a product of this context, of advocating that technoscientific decision-making processes should be opened up to involvement by laypeople, and understanding this will be important in seeking to account for why the literature relevant to RQ3 is so scarce.

Challenging the canonical view of science Until the 1970s, the dominant view in Europe and America was that the best way to promote the efficacy of decision-making processes when forming public policies was to found this process upon science (Brooks 1968), and that scientists should be afforded a great deal of freedom in defining the scope of, and in conducting, their research (Polanyi 1962; Price 1962, 1965). In particular, it was held that the scientific method was about the rational pursuit of objective truth and, hence, that it could determine which of the policy options available would be the best for society. Even in the light of a series of technoscientific disasters in

[7]Briefly, the term *'deficit model'* refers to a cognitive model of the public understanding of science that which explains any resistance or rejection by a layperson of advice or arguments put forward by a scientist in terms of a cognitive or knowledge deficit on the part of the layperson. It implies that the layperson's dissent would be allayed if only they were to be taught the necessary scientific facts and trained to think scientifically (see Miller 2001; Wynne 1991, 1992, 1996a, 1996b).

the 1960s and 1970s, which led to a fall in public confidence in science and technology, government programmes to restore trust (such as public education programmes or the setting up of public discussion groups about specific scientific issues) were founded upon these assumptions: as Nelkin (1979) reports in a study of such programmes relating to nuclear power, government initiatives were *'limited and controlled: information efforts were intended mainly to convince the public that nuclear energy is necessary and that it involves limited risk'*, and little, if any, tangible impact upon policy-making was seen (p. 117).

In the 1970s, a number of sociologists sought to challenge this view of science, their main aim being to promote substantive public participation in technoscientific decision-making processes. In the 1970s and 1980s, authors drew on insights from the emerging field of SSK and presented case studies of science-based policy making to argue for democratization in these processes. They argued that social and political factors had a fundamental impact upon the content of scientific knowledge and hence upon policy decisions that were legitimated by appeals to science (Gillespie *et al.* 1979; Nowotny 1979; Roy 1981), and hence that the decision-making processes that led to these policies should be democratized (Ravetz 1971; Noble 1977).

These calls for democratization, however, were not without their opponents: it was argued that such democratization would compromise the quality of these decision-making processes due to the apparent ignorance and inability of lay people to handle scientific knowledge (i.e. they were portrayed as irrational), (Roy 1981). In order to overcome such resistance, a more concerted and systematic programme to advocate democratization of these processes was undertaken by social scientists in the 1980s and 1990s, most notably by Brian Wynne, Alan Irwin and Sheila Jasanoff, amongst others, attempting to show that the apparent ignorance and irrationality displayed by laypeople was, instead, a result of alienation caused by the way technoscientists behaved when interacting with them (for instance, see Irwin 1995; Jasanoff 1997; Wynne 1982, 1991), and it is this body of work which will be discussed below.

Public alienation and the deficit model Within this body, there is a striking degree of unity regarding the methodologies they employ when conducting their empirical work, and the theoretical approaches which underpin their analyses, as well as in the conclusions they offer about how scientists perceive the lay public and how they seek to position themselves relative to the lay public. First, a brief overview of these common characteristics will be given, before a more detailed account of them is presented through a discussion of the particular case studies carried out within this paradigm.

These case studies share many methodological similarities. Most notably, they share an interpretive approach, which involves detailed qualitative study of specific instances where technoscientists and other scientific experts interact with lay people, typically involving a relatively small (typically, fewer than 50) group of people (MacIntyre 1995). Typically, this involves some mixture of: close observations of such interactions as they take place, for instance at public meetings where technoscientific issues are discussed (e.g. Kaminstein 1996; Roth *et al.* 2004; Wynne 1982, 1991) or of focus groups or similar gatherings (e.g. Davies & Burgess 2004; Yearley 2000); and detailed textual analysis of interviews and questionnaires conducted with those involved in such interactions (e.g. Callon & Rabeharisoa 2004; Irwin *et al.* 1996; Lee & Garvin 2003; Roth *et al.* 2004; Stilgoe 2007; Wynne 1991, 1996b). Furthermore, studies very often also involve close textual analysis of artifacts of such interactions, which may include: official reports produced in the cases where the interactions formed part of official government inquiries (e.g. Bauer *et al.* 2007; Stilgoe 2007; Wynne 1982, 2001, 2006), or archives or transcripts of various stages of these inquiries or other meetings between scientists and laypeople (e.g. Hinchliffe 2001; Kaminstein 1996; Roth *et al.* 2004; Stilgoe 2007); public statements made in the media (e.g. Jasanoff 1997; Roth *et al.* 2004); and expert advice given to laypeople, for instance in leaflets or via websites (e.g. Stilgoe 2007; Wynne 1996b).

The first feature of these case studies that shall be considered here are attempts to show that technoscientists and scientific experts (such as government advisors) continuously work to construct and maintain the terms upon which their interactions with laypeople take place, motivated by the goal of acquiring and maintaining their privileged positions within technoscientific decision-making processes (Wynne 2001). As Roth *et al.* (2004) argue, *'expert status'* should be understood as *'an outcome of social interactions rather than something that exists before and determines an interaction'* (p. 154), i.e. the categories of lay and expert need to be constructed. Thus, it is through their work that the technoscientists *'tacitly and furtively impose prescriptive models of the human and the social upon lay people'* (Wynne 1996a, p. 57), attempting to construct both their own identities and the identities of lay people in order to make it seem that laypeople *'are only capable of taking sentimental, emotional and intellectually vacuous positions'* whereas scientists deal with *'factual, objective and real knowledge'* (Wynne 2001, p. 445. See also Powell *et al*. 2011). Bound up in these constructed identities are notions of which roles can and cannot be legitimately performed by the various individuals during these interactions: when is it appropriate for an individual to speak or contribute, when should they remain silent, who is able to exercise authority over others and when they are able to do so, who should prevail when two or more individuals come into conflict, who has appropriate authority to determines who prevails, and so on.

Some of the moves made by technoscientists and scientific to frame and control the processes which occur during interactions between themselves and laypeople have been well-captured in a case study of a series of public meetings held by the US government agency the Environmental Protection Agency (EPA) for residents of an *'environmentally-threatened town'* in New Jersey, USA, regarding the proposed clean-up of a toxic waste dump located near their town (Kaminstein 1996, p. 458). These meetings were run by EPA-sponsored scientists, who performed a number of moves designed to create and maintain a boundary between the scientists and experts that would speak at the meeting, and the lay public attending the meeting, and to position the former group relative to the latter so that the views of the former would always take precedence over those of the latter. These moves included the physical layout of the meeting, where scientists were seated on raised platforms and behind tables, which served to differentiate the scientists from the public attending the meeting, giving the impression of the scientists having an elevated importance (for a discussion of the importance of physical positioning in structuring interactions, see Edelman 1985). Another move was structuring the protocols for the conduct of the meeting so that scientists could speak at any time they wished, for as long as they wished, and were able to interrupt contributions from members of the public, whereas the length of contributions from members of the public had to be kept within a certain time limit after which the contribution was cut-off (Kaminstein 1996. See also Roth *et al.* 2004).

On this view, not only did technoscientists carry out these moves in order to construct their own authority within interactions with laypeople, but, fundamentally, they also sought to impose a discourse in which scientific knowledge would be cast as objective and certain. For instance, Kaminstein (1996) reported how technoscientists sought to achieve this by praising contributions made by members of the public which seemed to employ this discourse, thereby encouraging the use of this discourse rather than other discourses.

Another example of how technoscientists attempted to impose such a discourse is given in Wynne's 1982 case study of the public inquiry held to decide whether to build a nuclear power station at Windscale in the UK, in which he was a participant-observer (as an advisor to certain pressure groups who opposed the power station), where those laypeople who sought to question the institutional commitments and trustworthiness of the technoscientists presenting arguments in favour of the proposed new power station were silenced by the chair of the Inquiry, who apparently vigorously ruled that such questioning was not a legitimate part of the Inquiry and that any attempts to ask such questions should be deleted from minutes and other official accounts of public meetings (Wynne 1982). Such moves can be seen as part of a strategy on the part of technoscientists to establish and maintain a boundary between themselves and dissenters from the lay population, a strategy which is well-documented in the literature and widely-held to be a fundamental and general feature of interactions between scientists and laypeople (Darier *et al.* 1999).

This attempt to establish the dominance of scientific discourse is echoed in findings from other case studies. The language used by technoscientists in the presentation of their arguments is often reported to be characterized by a lack of emotion and frequent use of technical terminology, for instance in the context of a public enquiry about a toxic waste dump (Kaminstein 1996), or in encounters between physicians and their patients (Lee & Garvin 2003). As a result, this means that testimony by laypeople that can be cast as non-scientific (such as knowledge claims that seem particular rather than universal and generalizable, or involving human emotion or values) are liable to be marginalized (Aiken 2009).

Furthermore, technoscientists often construct their knowledge claims in a way that conceals or downplays uncertainties in scientific knowledge, masking disagreement amongst scientists and thereby giving the impression of greater certitude than is perhaps actually the case (Lambert & Rose 1996; Wynne 1996a). What is presented by technoscientists during interactions with laypeople, then, are accounts which are purported to contain hard and indisputable scientific facts. Uncertainty and disagreements amongst scientists about these knowledge accounts, and the contexts in which they were generated, are not reported, resulting in the impression that these facts are universally true and therefore are directly applicable in the context in which they are being used (the interaction between the technoscientists and the laypeople). This, in turn, makes it difficult to dispute the appropriateness of deploying them in such interactions.

Using this type of discourse, then, makes it very difficult for a dissenter to challenge successfully the accounts provided by technoscientists: in the case of the Windscale Inquiry, for instance, the resultant assumption made by the chair of the Inquiry was that arguments made which opposed the accounts given by technoscientists were either factually inaccurate or simply irrational (Wynne 1982). The upshot, then, is that arguments voiced by laypeople are assessed against the statements of the technoscientists. Any dissent - which could take the form of directly challenging technoscientists at public meetings (Aiken 2009; Wynne 1982), refusing to believe scientific advice that mobile phones are safe to use (Stilgoe 2007), patient refusal to follow the advice of doctors (Lee & Garvin 2003), or other forms besides - is construed as subjective. This, in turn, is used to construct a model of the layperson as irrational, ignorant, and perhaps even *'intellectually vacuous'* (Wynne 2001, p. 445), i.e. the deficit model.

Indeed, it has been argued that technoscientists will persist with the deficit model, even when they face a great deal of pressure to open up their decision-making processes to laypeople, and will engage in strategies in such a way to deflect this pressure and maintain the position of science as the main knowledge-generating institution in society. In recent years, it has become increasingly difficult for technoscientists to simply dismiss the views of lay publics and cast them as having no value to such processes, in light of shifts in the political culture in the UK and the European Union in the past ten years in favour of greater democratization (for instance, see the report of the House of Lords Select Committee on Science and Technology (House of Lords 2000) in the case of the UK, and the report published by the European Commission called *Taking European Knowledge Society Seriously* (Wynne & Felt 2007)). Wynne reports that they have negotiated the difficult task of maintaining their authority and position within such processes, but also to demonstrate that they view public involvement as valuable by *'reinventing'* the basic deficit model to accommodate these competing aims (Wynne 2006, p. 214). For instance, in an analysis of reports produced by committees of government agencies dealing with regulation of genetic technologies (e.g. genetically modified food), Wynne shows how technoscientists engaged in sophisticated discursive moves to separate entirely the ethical dimension of regulatory decision-making from the scientific dimension, and then to define public concerns about genetic technologies as *'ethical rather than (scientifically) risk-focused'* (Wynne 2001, p. 447.). This allows the impression to be given that public views are given *'autonomous policy weight'* (*Ibid.*), with their views being incorporated into ethical considerations, but still maintains the privileged position of the technoscientists as being the sole providers of objective arguments to determine the final outcomes of decision-making processes (see also Felt *et al*. 2009; Bauer *et al.* 2007; Parry 2009).

These case studies together suggest, then, the extent to which technoscientists will engage in strategies to resist opening-up scientific knowledge-generating processes to laypeople. They will seek to structure interactions with laypeople according to technoscientific discourses, against which any dissent on the part of laypeople is contrasted as irrational, which in turn enables the exclusion of laypeople from making a substantive contribution to technoscientific decision-making processes. Furthermore, when faced with increasing pressure to allow laypeople's perspectives to be included, technoscientists have been seen to engage in further boundary work to find ways to cast these perspectives so that they give the appearance of seeing them as valuable whilst nevertheless maintaining the privileged position afforded to technoscientific discourse.

Lay knowledges Another component of these case studies is their attempts to demonstrate that this deficit model does not capture the way in which laypeople actually handle and interpret knowledge claims and advice presented by technoscientific experts. Rather than dissent or resistance on the part of laypeople being a result of ignorance or irrationality on the part of *'intellectually vacuous'* individuals (Wynne 2001, p. 445), these case studies attempted to show that such a response to science is actually a product of sophisticated moves and evaluations on the part of these laypeople. This aspect of the literature is indeed relevant to answering RQ1 and RQ2 for two, related reasons. The first is that the failure of the deficit model to truly capture how laypeople handle and evaluate scientific knowledge results in the alienation of laypeople, which clearly suggests that technoscientists involved in Citizen Cyberscience Projects which have successfully engaged volunteers may well seek to construct their interactions with these volunteers in ways that differ significantly from the deficit model. The second is that the analyses in these case studies lend themselves to conclusions regarding how technoscientists might interact with laypeople in a way that does not alienate them. These should aid in understanding the instances where technoscientists and certain groups of laypeople have managed to engage each other, which are presented in subsequent subsections of this Chapter, and should also inform answers to RQ1 and RQ2 when considering how technoscientists in CCPs have sought to engage and retain volunteers.

These case studies draw on the conclusions from STS that the process of knowledge production is contingent, and they seek to show that the way by which laypeople attempt to make sense of the world, as with scientists, involves a complex interplay of various social, cultural and material/physical factors. For instance, just as Knorr-Cetina (1999) demonstrated that different scientific disciplines could be seen to involve different and frequently incommensurable epistemic cultures, so too are laypeople members of epistemic cultures, in which knowledge claims and advice presented by technoscientists are filtered, assessed, and interpreted according to cultural norms and traditions which may be particular to the specific communities and groups (geographical, social etc.), as well as knowledge shared by members of the community about the particular community and its surroundings, of which the individual layperson is a member (Cvetkovich & Lofstedt 1999). One example of such knowledge is given by Wynne in a case study of Cumbrian sheep farmers: they have a very sophisticated knowledge about local soil types, the farming industry, local climates and so on. Wynne coins the term *'local knowledges'* to capture this (Wynne 1996a, p. 44. A related concept is that of *'situated knowledges'*, introduced by Haraway (1988)). Additionally, people will also measure *'expert knowledge...against elements of their own already-tested knowledge and direct experience'* (Wynne 1991, p. 115. See also Connor & Siegrist 2010; Davies & Burgess 2004). Finally, just as case studies in SSK (Sociology of Scientific Knowledge) suggested that, at least in part, scientists might choose between competing knowledge claims on the basis of which knowledge claim best supports the pursuit of their own interests within the social orders that they inhabit, so laypeople, too, are driven by the protection and advancement of their own interests when making choices about whether to reject or accept the knowledge claims presented by technoscientists (Kahan *et al.* 2011; Wynne 1982). They will also tend to privilege knowledge claims that help to maintain the social order of which they are part, in other words those claims which are consistent with the order's shared norms and values, and common knowledge, and will tend to ignore or reject those claims that are in conflict (Carlisle *et al*. 2009; Ho *et al*. 2010). All of this taken together - epistemic cultural norms, consideration of personal interests and experiences and of the interests of the social orders in which they are embedded - will determine whether a layperson accepts or dissents from the advice of technoscientists (Wynne 1991).

To better illuminate what this means, attention will be turned to what is perhaps considered to be the exemplary case study on this subject. Wynne (1996a, 1996b) conducted a study of sheep farmers in Cumbria, UK, in the wake of the Chernobyl nuclear disaster of 1985. Tests of the local soils in 1986 revealed dangerously high levels of radiation, and the response of the government was to place temporary restrictions on the movement and selling of sheep farmed in this area. Scientists from the government were dispatched to Cumbria to conduct various tests and devise a strategy involving giving advice to farmers regarding where they should or should not allow their sheep to graze and for how long, with the apparent ultimate goal of lifting the restrictions as soon as possible. After a brief period of time, however, the farmers began to display distrust of the advice given by, and resistance to cooperating with, the government scientists.

Wynne accounts for this by arguing that the farmers rejected the framing of the scientists' advice as being based purely upon the application of scientific knowledge which was certain, objective and universal; the farmers were aware of the contingent nature of the processes in which this knowledge was produced. For instance, their prior experiences regarding the institutions represented by the scientists led them to suspect that the scientists' blaming of the contamination upon the Chernobyl disaster was, in fact, a cover-up, hence provoking mistrust of the scientists' advice: even before the Chernobyl disaster, the farmers had long suspected that the nearby Sellafield nuclear power station, in whose success the government had a substantial political investment, was contaminating their soil and this suspicion was further aroused by the institutional secretiveness they had encountered when they asked the relevant government ministry whether this was the case (Wynne 1996a). Similar observations regarding laypeople's mistrust of expert advice due to prior knowledge of these experts' institutional commitments are made by Allen (2007) to account for why residents of New Orleans who had been displaced by Hurricane Katrina in summer 2005 did not trust reassurances from public health officials that it was safe to return to their homes: these residents had previously observed instances where these officials had refused to address public concerns about pollution, instead being seen to protect polluting industries. Irwin *et al.* (1996) also echo this in a study of residents in the Greater Manchester area of evaluations of assurances from scientists in the chemical industry that local chemical plants did not pose a threat to public health.

Returning to Wynne's case study of the Cumbrian sheep farmers, other factors were identified which led the farmers to reject the dominant technoscientific framing of the advice and arguments presented by the scientists were also identified. These include the farmers' recognition of the uncertainties lying behind the advice, which was made visible when the scientists abruptly changed their advice (for instance, in extending the length of the ban on the sale of the sheep, and changing the areas where the farmers are advised to graze their sheep), and the scientists' subsequent refusal to acknowledge the existence of these uncertainties provoked a further loss of trust on the part of the farmers (see also Darier *et al.* 1999).

Additionally, the sheep farmers also drew on their local knowledge regarding climates and soil patterns and of the particularities of the livestock trade to evaluate the assumptions made by the government scientists when they were formulating their models and advice. For instance, the models used by the scientists to estimate how long it would be before the contamination was washed out of the soil were rejected by the farmers because they involved certain simplifying assumptions about the type of soil to be found on the affected land and about the local climate: the farmers knew that the patterns of soil types was substantially more complex than was implied in the models, and that predicting when contamination would be washed from these soils was further complicated because the properties of these types of soil would change depending on local weather conditions. That laypeople will reject scientific arguments and advice because they believe, based on their own knowledge, that the assumptions underlying these arguments are false, has been echoed elsewhere. For instance, medical practitioners often objectify the body of a patient, and their process of making diagnoses and devising courses of treatment or formulating lifestyle advice is founded upon a model of an idealized notion of a typical patient, rather than acknowledging the particularities of each individual. A patient then evaluates this advice according to their particular knowledge of their own personal and family's medical history and how proposed treatments or courses of action would impact upon their own lives (for instance, in the case of a cancer patient who rejects a grueling course of treatment with only a slim chance of success in order to some time with their family while not being severely impaired by treatment side effects (Lambert & Rose 1996; Lee & Garvin 2003)).

A final insight from Wynne's case study demonstrates that the course of action taken by a layperson in response to the arguments given by technoscientists will also be determined by judgments regarding how best to protect their interests within the social order, including the protection of the networks in which they are embedded. It was this which Wynne uses to explain why a substantial minority of those interviewed accepted, rather than rejected, the advice of the government scientists. These individuals were found to have close family members and neighbours whose livelihood depended upon the continued running of the Sellafield station: to reject the advice of scientists was to endorse the argument that Sellafield was responsible for the unsafe levels of radiation found in the soil which, in turn, would threaten the livelihoods of those who worked there if Sellafield was deemed a risk to health and operations.

The literature presented so far displays a remarkable degree of consensus around a few key ideas. It is argued that technoscientists engage in two moves when they encounter the public, constructing both their own positions and identities, and those of lay people. They present accounts which conceal the contingent nature of the processes by which they arrive at their knowledge, instead presenting their claims as completely objective and their knowledge as certain. Against this, laypeople who dissent or show resistance to the technoscientists are cast as irrational or ignorant (the deficit model). Elsewhere in the literature it is argued that the deficit model does not capture the true nature of how laypeople understand accounts and advice presented to them by technoscientists and scientific experts. Instead, laypeople are members of communities which possess complex epistemic cultures. They: 1) display awareness of the contextual and uncertain nature of the knowledge underpinning these accounts, and of the institutional and other social commitments of these technoscientists and will make assessments of the trustworthiness of these technoscientists accordingly; 2) possess their own local knowledge against which they will measure the knowledge claims advanced by the technoscientists; and 3) they will seek out information from other sources; and they will more readily accept these claims if they believe that their widespread acceptance will serve in their own interests.

From this, it has been argued (as suggested above) that the way in which technoscientists construct their interactions with the public has can alienate the lay public. A refusal to acknowledge the contextual and contingent nature of their knowledge in the case study can result in a loss of credibility and trustworthiness on the part of scientists in the eyes of the public, leading laypeople to disregard advice from, or reject arguments advanced by, the technoscientists. A failure to acknowledge the value of lay perspectives can lead to laypeople feeling patronized and undervalued, and believing that the expert discourses do not capture the full complexities of the situation, resulting in *'backlash in the form of cultural alienation'*, (Wynne 1993, p. 323. See also Kaminstein 1996; Powell *et al.* 2011; Wynne 1991, 2006).

Reported instances of alienation include: the rejection by some patients of doctors' diagnoses of them as ill (Lambert & Rose 1996), whereas some others might accept the diagnosis but will subsequently opt out of the healthcare system altogether (Callon & Rabeharisoa 2004; Lee & Garvin 2003); *'public protest and withdrawal* *of legitimacy from...nuclear power'* (Wynne 2001, p. 450; also Wynne 1982); a collapse of trust in government scientists and the institutions during the BSE crisis in Britain in 1996, so that *'people looked to other institutions - the high street butcher, the restaurant, the media, the supermarket - for information'* (Jasanoff 1997, p. 223); and anger at feeling patronized by scientists when, during the course of a public meeting about the proposed siting of a toxic waste plant, scientists would offer comments on whether they believed a question asked by a member of the public was a good question (Kaminstein 1996).

Social learning Given the diagnosis outlined above regarding why laypeople might become alienated from the institutions of science, the prescription for re-engaging the public seems to be what Wynne calls a process of *'social learning'* on the part of the scientists (Wynne 1992, p. 282). Such a process would involve, firstly, *'reflexivity'*, which Wynne defines as *'the process of identifying, and critically examining...the basic, preanalytic assumptions that frame knowledge-commitments'*, in other words, an acknowledgment by technoscientists of the contingent and contextual nature of the way in which their knowledge is generated (Wynne 1993, p. 324). Social learning would also involve the identification and appreciation of *'other actors' (i.e. laypeople's) precommitments and framing processes'*, in other words, recognizing and understanding how laypeople evaluate and interpret the scientific knowledge that is presented to them in terms of their own epistemic cultures, including other sources of knowledge (Wynne 1992, p. 282).

The insights presented above are very promising in the context of answering Research Questions 1 and 2 (RQ1 and RQ2) as scientist/laypeople interactions are at the very heart of CCPs. However, as shall be explained below, the context of CCPs are also interestingly different from what has been considered in the literature presented above. This literature therefore needs extension to understand what occurs in CCPs. I hope to develop this fresh perspective in the rest of this thesis.

RQ1 deals with scientists and software engineers who have had no experience of involvement with CCPs. It is possible that such individuals might, as they begin to embark upon their involvement with CCPs, seek to construct their interactions with volunteers in a manner similar to that described in the deficit model. When entering into their interactions with laypeople in CCPs, they will be emerging from institutional backgrounds where they have been thoroughly embedded in technoscientific culture. Furthermore, it is highly likely they will lack prior experience of situations where they have had to undergo or display social learning of lay epistemic cultures and laypeople's knowledge. It might, therefore, seem plausible that such scientists' and software engineers' perceptions of the project volunteers will concur with the deficit model.

However, in terms of RQ2, it seems unlikely that project team members who have already been involved with CCPs would support a project in which the dominant construction of the project volunteers is similar to that described in the deficit model. This is because these projects simply cannot afford to alienate project volunteers. The conclusions of the research presented above suggest, then, that attempting to construct their interactions with potential volunteers in such a way that accords with the deficit model is likely to prove a disastrous strategy for those involved in setting up and running a CCP, alienating potential and existing volunteers. Therefore, if Wynne's diagnosis of, and prescription for, the processes by which laypeople become alienated from scientific institutions is largely correct, then it may be reasonable to presume instead that a CCP's scientists and software engineers will display evidence of having undergone a process of social learning. Literature that deals with situations where one might assume that technoscientists will have undergone a process of social learning, and will be displaying this in their interactions with laypeople (i.e. situations where the balance of power between technoscientists and laypeople is far more in favour of the latter than in the case studies presented above) will form part of the next subsection, which considers relationships between technoscientists and laypeople which have not resulted in the alienation of the latter.

2.2.2 Alternatives to the deficit model

According to the literature reviewed so far, then, the deficit model seems to capture the dominant view amongst scientists regarding non-scientists and, furthermore, that it is in the scientists' interests to ensure that the boundary implied in this model is enacted and defended in public, and they will diligently engage in strategic moves to ensure this. This subsection considers examples in the literature where this boundary has been crossed or dissolved, and is split into two parts. Those examples in the first part are characterized by the maintenance of this boundary, with certain laypeople being allowed to cross this boundary by adopting the same discourses as the technoscientists; those in the second part feature the dissolution of this boundary, with technoscientists and their representatives displaying a substantial amount of social learning about the epistemic cultures of laypeople and reflexivity regarding their own knowledge-generating processes.

Examples of laypeople crossing the boundary The case studies presented here purport to show instances where scientists have initially been resistant to the inclusion of laypeople but have subsequently changed their views of some of these laypeople - namely, those who adopt the discourses of technoscience - with the result that these technoscientists have ultimately allowed these particular laypeople to have a substantive impact upon their knowledge-producing practices. The key feature underpinning these case studies is that they involve issues where the stakes are extremely high for the lay people involved, which provides a very strong motivation for them to engage in social learning regarding, for instance, the discourses of science; by contrast, the scientists have little incentive to engage in social learning about the lay people involved.

Initial focus will be on one such case: the key features will be drawn out and it will then be explained how these features are common to other accounts in the literature of similar instances. This case is Epstein's studies of AIDS treatment activists in the USA in the 1980s, where some lay activists were successful in persuading medical researchers and government agencies regulating drugs trials to change some of the protocols regarding the conduct of clinical trials of new drugs for HIV/AIDS (Epstein 1996. See also Epstein 1995, 2000). Many sufferers of HIV/AIDS, and treatment activist groups, desperately needed access to new drugs in the late 1980s and very often participation in clinical trials was the only option for gaining access to these drugs. Their desperation for new drugs stemmed from the fact that HIV/AIDS had spread very rapidly in the early 1980s amongst particular sections of the population (gay people, intravenous drug users), but that initially, there was no effective treatment available until drugs began to be developed in the mid-1980s. At the time, these new drugs had to go through a lengthy series of trials conducted according to a set of norms laid down by US government agencies before they could be legally prescribed by physicians, norms which restricted access to these drugs and meant almost certain death for those not given the drugs. For instance, one such norm was double-blindness, where half of participants in a trial are randomly given the new trial drug and the rest are given the placebo treatment (usually the existing standard treatment): those assigned the placebo faced certain death. This, combined with other features of the drug-trialing process, such as the length of time it took for a drug to get approval when large numbers of people were dying, proved unacceptable to many, who formed activist groups to advocate for change in the trial process.

In the late 1980s, the strategy adopted by some of these activists was to learn the discourse of medical science: they invested substantial amounts of time to learning medical terminology and reading scientific journals, and learnt how to critique scientific claims in the same terms as scientists, acquiring what Shapin (1990) termed *'cultural competence in science'* (p. 993). The scientists, including those with highly influential positions in government drug licensing bodies, were, following initial scepticism, reportedly deeply impressed by these activists. They increasingly found their knowledge regarding the needs of patients and how clinical trials operated in reality, often grounded in personal experience or the experiences of others, useful in the design of clinical trials. By 1990, many of the norms regarding clinical trials, such as double-blindness, had changed as a result of the advocacy of these activists, who were very often invited to join committees designing and regulating trials for new drugs. At the same time, however, the research scientists still moved to exclude those activists who had not adopted scientific discourses, which suggests that these research scientists were still very concerned with maintaining a cultural boundary to exclude those who were not deemed to be culturally competent in the relevant medical science (Epstein 1996).

The features highlighted above are common to other accounts where non-scientists have gained access to, and substantive influence in, scientific knowledge- and decision-making processes. One key feature is that non-scientists have sought such involvement in response to situations where serious issues of health are at stake. These include: seeking to influence public policy regarding the scope and direction of research in genetics, and designing of genetic counselling services (where, for instance, individuals are tested to see whether they carry genes causing particular diseases) (Kerr *et al.* 2007); and deciding the allocation of public funding between different areas of research regarding the prevented and treatment of breast cancer (for instance, whether funding should be allocated towards researching the environmental causes of cancer, cures, and palliative care) and the standards by which new treatments are assessed (for instance, the extent to which a treatment should be judged on whether it provides an outright cure, or whether major consideration should also be made of quality of life issues relating to treatment side- effects) (McCormick *et al.* 2003).

The other main feature common to these accounts is that scientists were initially very reluctant to open up their work to non-scientists (this is implicit in most accounts, but is made more explicit in Kerr *et al.* 2007. Parthasarathy (2010) calls this the *'expertise barrier'* (p. 355)), and that this resistance was overcome only after a concerted effort on the part of these laypeople to learn the cultural norms of science, in effect becoming *'lay-expert...hybrids'* (Kerr *et al.* 2007, p. 407), or *'protoprofessionals'* (Nowotny 1993, p. 313).

The link between the extent to which these non-scientists behave according to these norms and the extent to which scientists have allowed them to have an impact on scientific decision-making processes is made particularly explicit in Barbot & Dodier's study of decisions regarding the particular drugs trials to which HIV/AIDS patients would be allocated: in the case of those patients who displayed ignorance of the basic science relating to HIV/AIDS and the trials, the clinicians took complete control of this decision; those patients who learnt the protocols of various clinical trials were granted a certain amount of autonomy over their choice of which trials to enter, but the clinician still made the final decision; finally, those who not only learnt about these technicalities but also displayed critical evaluation of the credibility of various sources of information were given almost total autonomy over the choice of clinical trial (Barbot & Dodier 2002).

Here, some literature has been presented which deals with situations where relationships have been established and maintained between laypeople and scientists which have resulted in the former having an impact upon technoscientific decision-making processes. In these cases, there was reluctance on the part of scientists to engage with the laypeople; it was persistence on the part of certain laypeople, which involved concerted efforts to become culturally competent in the relevant areas of science, which enabled them to gain access to the scientists. At first sight, then, these situations may seem to have little to offer in terms of answering RQs 1 and 2. After all, there is a great deal of interest on the part of scientists in CCPs to engage with the public; certainly, there is no expectation that the volunteers should learn scientific discourse in order to be allowed to participate. However, as discussed in subsection 1.1.3, volunteers within a CCP might be pursuing a range of interests, for which they may believe they need to enroll the CCP's scientists or software engineers and may therefore adopt the discourses of these groups in order to do so. For instance, they might wish to be appointed as moderators in the online forums, or seek to gain influence in how the project is designed and developed.

Engaging laypeople So far, it seems that in the literature that there is a very strong consensus that the deficit model (broadly defined) is a good description of how scientists seek to construct and maintain their relationships with members of the public. In this sense, it can be seen that a failure by a technoscientist to display social learning can be seen as part of a performance of boundary work to demarcate science and scientists from non-science and non-scientists. Even in the cases of the examples presented in the previous subsection where substantive cooperation has occurred between scientists and non-scientists, there has been evidence of social learning displayed on the part of scientists. Instead, they have involved the adaptation of non-scientists to a particular scientific culture, rather than the other way round. There seems, therefore, to be a great deal of resistance on the part of technoscientists to displaying social learning in their interactions with laypeople as to do so would be to dissolve boundaries between scientists and non-scientists, thereby jeopardizing science's position as the pre-eminent knowledge generating institution in society.

However, as noted previously, it is very much in the interests of scientists and software engineers in a CCP to engage members of the public, and it therefore seems likely that the pursuit of this interest would induce these particular technoscientists to display social learning, rather than constructing boundaries between themselves and the public by constructing relationships with members of the public along the lines of the deficit model. Therefore, we need to turn our attention to the literature which covers situations where individuals such as technoscientists or science journalists interact with members of the public regarding technoscientific issues and where it seems to be very much in the interests of these individuals to display social learning. Such examples might include: journalists working in commercial news outlets, whose success is determined by the ability to attract consumers from amongst the general public; popular science publications (books and magazines), who likewise need to attract consumers; and museums, who need to attract visitors.

Focusing first on science popularizers (journalists, museum curators, writers of popular science), there is substantial evidence in the literature that these communicators, as expected, display a very sophisticated level of social learning in their interactions with members of the public, for instance in the articles they write or museum exhibits they design. In study after study in the literature, it is argued that the essential features of the way in which science popularizers construct their accounts of technoscience cannot be captured by what is known as *'the canonical model'* (Bucchi 1998, p. 3, also known as the *'diffusion model'* (Lewenstein 1995, p. 348)), which as shall be explained below, is analogous to the deficit model and hence would entail no display of social learning on the part of science popularizers; instead, it has been argued, they proceed in a fashion that is captured by *'the contextual model'*, whereby they display a highly sophisticated degree of social learning, as will be explained below (Lewenstein 1995, p. 350).

According to the canonical model, scientific knowledge is communicated in a one-way direction, firstly from scientists to the media, and then from the media to the public, with the modifications at each stage of communication merely being a case of simplifying the content of the scientific statements (Hilgartner 1990), with the supposition that this will result in a situation where *'lay audiences simply absorb, in an impoverished and lessened form, ideas which stem from scientific activity'* (Bucchi 1998, p. 5). The parallels between this canonical model and the deficit model are clear: scientific knowledge, as it proceeds from scientists to the public, is cut free from the context in which it was generated, so that the social and institutional commitments of those who produced it are deleted from the accounts presented to the public. Furthermore, it is expected that it will simply be accepted by the public: there is no acknowledgment of the complexities of the processes by which the public evaluate such accounts of knowledge. Hence, the canonical model assumes no display of social learning on the part of the science popularizers.

By contrast, the contextual model presents science communication to the public as *'a complicated tangle of processes and transformations through which science is appropriated, used or simply neglected by different audiences'* in accordance with the pursuit of interests by those involved (Bucchi 1998, p. 7). In support of such a model, many analyses of items of scientific communication have been conducted, and these suggest a high level of sophistication on the part of science communicators in terms of understanding the lay epistemic cultures and the contexts in which their accounts will be received. This is because popular accounts of science are often shaped by the need to engage members of the public, for instance when commercial interests are involved or museums must justify their requests for public money. To achieve this, a science popularizer will very often need to produce accounts of science which members believe are relevant to their lives in terms of how it might mitigate threats to their lives (e.g. to avert natural disasters) or result in new risks (e.g. environmental catastrophe) or provide opportunities for improved lives (Beck 1992)[8], but they will also need to provide accounts that can stand up to evaluation by members of the public according to standards of their own epistemic culture, as well as against personal knowledge and experiences and whether the knowledge accounts support jeopardize the pursuit of their interests within the social orders they inhabit (see subsection 2.2.1).

In other words, scientific popularizers, such as journalists, appear to display a great deal of social learning about their audiences. One aspect of this is that they seem aware of what members of their audience seem to particularly value in their everyday lives. As Gross & Lewenstein put it, *'stories about science in the mainstream press tend to fit into two simplistic formats:* "Gee whiz, isn't science interesting and useful" *or* "Beware of toxic chemicals lurking in your backyards"*'* (Gross & Lewenstein 2005, p. 9). Journalists' selection of which scientific issues they report, and the manner in which they construct their narratives about these issues, is based upon what they believe can be portrayed as a threat or a source of sustenance for these relationships (Einsiedel 1992). For instance, widespread coverage of AIDS in the media only began when it became clear that a large number of heterosexuals, as well as homosexuals, were being infected by the HIV virus, and hence could be cast as relevant to the public-at-large and portrayed as a threat to individuals' network of social relations in the sense that anybody could be at risk (Nelkin & Gilman 1988). Other reported examples include: the portrayal of climate change as transforming the immediate physical surroundings of the target audience (for instance, by pictorial displays of familiar landmarks being submerged, or transformed into tropical paradises) (Weingart *et al.* 2000); how nanotechnology might affect human bodies (Gross & Lewenstein 2005); the economic benefits promised by biotechnology (Nisbet & Lewenstein 2002); and the coverage of genetics in terms of whether it might offer the prospect of identifying violent criminals before they are able to commit crime (Nelkin 1994).

In the examples given above, it can be seen that a great deal of social learning appears to be displayed by these science popularizers, and they are thus able to present science in a way that members of the public find

[8] I accept that many members of the public are curious about matters of science which does not seem to relate directly to their own everyday lives, and will consume popular accounts of these matters. However, the number of such accounts in the media is dwarfed by the number of stories involving reporting of technoscience which laypeople deem relevant to their lives (for instance, see Gross & Lewenstein 2005; Nelkin 1995).

comprehensible and relevant. That this happens could provide a useful insight for understanding how scientists and software engineers in CCPs seek to structure their interactions and relationships with project volunteers, given that a primary challenge for them is to engage members of the public.

It is, however, is questionable to what extent a CCP's scientists and software engineers might display a matching level of sophistication in terms of social learning. The reason for this can be seen by noting the differences in the backgrounds of those involved in running CCPs, and science popularizers, particularly focusing on the cultural differences between their institutions and training. For instance, a science journalist will have undergone a period of formal training and apprenticeship within journalism, during which they will have been thoroughly enculturated: both journalism as a profession and the individual newspaper will have a culture whose norms largely determine how members of that culture (i.e. the journalist) should conduct their work, including: how topics and events are assessed in terms of whether they are newsworthy; decisions about which features of the story should be particularly highlighted and bow these features should be related to each other, such as which members of the population should be portrayed as being especially affected by the story; whose opinions should be canvassed and reported; and the language and style used in writing the story (Bucchi 1998). In other words, the content of news reports about science can be seen as being determined by such cultural norms. Hence, any social learning displayed by a science journalist should be seen not as a product of social learning on the part of that individual journalist alone, but rather as a result of the operation of cultural norms which are the result of social learning accumulated over many years by various members of the institution all responding to the need to engage members of the public. In contrast to this, scientists' training and apprenticeships generally take place in institutions whose success is not dependent upon direct engagement of the lay public and, indeed, where many of the cultural norms regarding interactions with members of the public appear to direct its members to perform boundary work (Gieryn 1983, 1999).

The question, then, is whether any literature exists which offers accounts where scientists appear to display social learning about lay epistemic cultures, particularly cases where they encounter situations requiring them to engage members of the public for the first time? Despite a thorough search of the literature, only one account has been found which seems promising in this respect (Hartman 1997). It details the experiences of scientists in an organization, *Earthwatch*, founded in 1972. It runs scientific research projects monitoring environmental phenomena around the world which involve both scientists and members of the public, the latter having paid a participation fee to fund the project. This means that the scientists are financially dependent upon the project. The scientists display a broad range of attitudes towards the participants: some see them merely as sources of labour, capable only of carrying out tedious or routine work, such as collecting fossils, or see the purpose of their involvement mainly in terms of having memorable experiences, including social contact with members of the local culture - in both cases, suggesting a deficit model; others, by contrast, see the lay participants as possibly making substantive contributions to the scientific knowledge produced by the project, and assign them roles where they are able to do so - a display of social learning. Unfortunately, however, these differences are not really accounted for by the author. Indeed, to the extent that they are, they are portrayed as being the attitudes held by the scientists before they joined Earthwatch, and it is therefore difficult to see how any social learning may have come about during the course of the scientists' interactions with the public.

So far, the literature relating to interactions of laypeople with technoscientists and individuals who seek to engage members of the public technoscience has been reviewed, and, as has been indicated throughout the review, the insights from this body of literature promises to help inform answers to RQs 1 and 2. However, it seems that this literature has little to offer answers to RQ 3, which is about how certain groups of volunteers within CCPs seek to structure their relationships with other volunteers, and the (lack of) literature relevant to answering this question will be considered in the next Section.

2.3 Interactions amongst laypeople

A very thorough search of literature has been carried out, but has resulted in failure to find any accounts of interactions between laypeople or lay groups about issues relating to technoscience. This short Section identifies the main bodies of literature where such accounts might have been supposed to be located, and briefly seeks to account for why they are absent.

One area of literature where relevant work might have been expected to have been found is some of that considered in subsection 2.2.1, namely case studies of the dynamics of technoscientific controversies and decision-making processes which were used as the basis of critiques of technoscientists' attitudes towards, and treatment of, lay people and their knowledge. Given the diversity of viewpoints held by laypeople, and the range of (potentially conflicting) goals they possess, it might have been expected that the dynamics of how various non-scientists and lay groups interactions with each other would provide rich source of material for the authors of these studies. For instance, they might have studied how various non-scientists sought to cast other non-scientists as possessing or lacking the legitimacy to participate in the various processes, or how lay groups reveal their views of how other lay groups interpret and handle knowledge. Indeed, Powell *et al*. (2011) acknowledge that relational dynamics and interactions amongst lay participants can impact upon their opinions as much as, if not more, than their interactions with scientific experts.

Nevertheless, this does not seem to have occurred: such case studies focused on the dynamics of interactions between scientists and the public; less so on interactions between various lay groups. To understand why, it is necessary to revisit the underlying agenda motivating the authors of these case studies: namely, to open up technoscientific decision-making processes so that members of the lay public might have a substantive impact upon these processes through arguing that these processes are impoverished by the exclusion of lay perspectives (MacNaghten *et al.* 2005; Roth *et al.* 2004). It is easy to see how highlighting any conflict arising between different members of the public might have proven counter-productive to these authors' agendas, in the sense that messy squabbling and status-grabbing between various lay people is likely to make lay participation seem less attractive, both to the scientists who need to be persuaded to open up their knowledge- and decision-making processes, and to institutions such as governments, to whom the apparent efficiency and rationality promised by scientific-based systems of policy-making is appealing (Nowotny 1981). It is not surprising, then, if, in their case studies, the authors who wished to promote public participation chose not to highlight frictions and debates between lay groups and individuals.

There exist in the literature case studies of other situations involving debates and discussions about issues of technoscience, studies which might be expected to contain accounts of interactions between various laypeople and lay groups. One group of situations of which there are a particularly high number of case studies involve what are known as *'public engagement mechanisms'*, events organized by governments or other public bodies where groups of non-scientists - often with some input from technoscientists - are gathered together to discuss and arrive at conclusions regarding a particular aspect of technoscientific policy-making (Lengwiler 2008). There have been a plethora of these in light of a political culture in Europe which increasingly favours public participation in technoscientific decision-making processes (Wynne 2006), and it might be supposed that they offer a particularly promising site of research for studying the dynamics of debates amongst non-scientists and lay groups about technoscientific issues. Certainly, a number of evaluations of specific examples of participation mechanisms have been conducted by social scientists with a normative commitment to deepening the scope of the impact that lay people have on technoscientific policy-making (for an overview of many of these, see Rowe & Frewer 2004), and it might therefore be thought that these studies would be concerned, at least in part, with how the issues are discussed and contested amongst the participants in order to ensure that all participants have a fair chance to contribute.

However, this does not appear to have been the case and is very difficult indeed to find any descriptions or perspectives on how discussions unfolded, how disputes and conflicts between participants emerged, were negotiated (including how the status of various participants and their legitimacy to make particular contributions were contested), and were ultimately brought to closure. This might in part be due to difficulties encountered by those studying public engagement mechanisms in terms of gaining permission from the organizers to actually observe these discussions: many of these mechanisms involve small- group discussions (typically between five and 15 participants), and the organizers often felt that observers might inhibit participants from expressing themselves fully (Barnes 1999; Rowe & Frewer 2004). However, even those studies which did involve observations of all stages of a mechanism either tend not to present any accounts of the discussions themselves (for instance, see Joss 1995; Lee 1995), or, in the rare cases when they do, studies report the dynamics of interactions taking place between lay participants and the scientific experts, rather than discussions amongst lay participants (Bennett & Smith 2007; French & Bayley 2011). Instead, they focused upon asking the lay participants' perceptions of their interactions with the technoscientists and other officials involved (Rowe *et al.* 2004. See also Joss 1995; Plumlee *et al.* 1985; Walls *et al*. 2011), or whether the conclusions arrived at by the participants did indeed have a substantive impact upon the policymaking process (Einsiedale *et al.* 2001; Guston 1999; Walls *et al*. 2011), or the impact of the method of communication amongst participants (e.g. an Information Communication Technology such as Skype) on their ability to express themselves and understand others (Delborne *et al.* 2011), or evaluating the extent of any such impacts against the expense of running the mechanism in order to assess cost-effectiveness (Kathlene & Martin 1991; Petts 1995). In short, then, there was little coverage of the dynamics of discussions between the lay participants themselves (it should be noted here that French & Bayley (2011) did set out to do this in one of the case studies they present, but they ultimately found that they *'could not accumulate sufficient evidence to support firm conclusions'* (p. 251)).

2.4 Conclusion

Broadly speaking, the relationships between technoscientists and laypeople presented in this literature can be separated into three categories, and the nature that these relationships take can be seen as being determined by the interests of the parties involved. One category concerns situations where technoscientists construct their interactions with laypeople along the lines of some form of the deficit model, leading to the alienation of the laypeople involved because, it is argued, such constructors acknowledge neither the contingent and contextual processes by which scientific knowledge is generated nor the real nature of the complex processes by which laypeople evaluate this knowledge. In these cases, the technoscientists are performing boundary work in order to promote and maintain the interests of science as an institution, and any resultant alienation of the laypeople with whom they are interacting is a price worth paying for this. The second category involve situations where certain lay members of the public or lay groups have been fully-integrated into, and make substantive impacts upon, at least some stages of technoscientific decision- making processes. Here, it appears that some technoscientists have allowed the lay-science boundary to be breached, but in reality, the members of the public allowed to breach this boundary have adopted the discourses of science This apparently occurs when certain lay people feel a desperate need to impact upon these processes, for instance in the case of Epstein's HIV/AIDS patients, and are thus willing to commit to these discourses fluently in order to engage the technoscientists (Epstein 1996). The final broad category involve situations where the situation is reversed: it is the members of the public who must be engaged, with the result that those who wish to engage the public (such as science journalists) must display social learning regarding the epistemic cultures, knowledges, and social interests of laypeople.

Insights from these three categories could prove useful in answering RQ1 and RQ2, regarding how scientists and software engineers involved in CCPs attempt to construct their relationships with project volunteers, although it seems that, at best, they can offer only partial answers. The first category of relationships presented in the literature seem to emphasize the extent to which technoscientists seek to create and maintain a boundary between themselves and laypeople, which should be borne in mind when considering how scientists and software engineers behave in the context of CCPs given that their backgrounds are in institutions with a culture of the performance of boundary work (Gieryn 1983, 1999), but the literature suggests that this might not be a sustainable strategy for CCPs to follow given that it often leads to the alienation of laypeople. The third category, involving people who seek to engage members of the public about issues related to technoscience, suggests that a strategy adopted by scientists and software engineers involved in CCPs could be to display social learning in some way, although the extent to which they mould their discourse around the contours of lay epistemic culture remains to be seen. Finally, the second category appears to offer little in terms of answering RQ1 and RQ2, given that they involve situations where there was little pressure on technoscientists to engage laypeople, and hence that laypeople needed to adopt the discourses of science to be able to engage and enroll the technoscientists. However, as noted previously, volunteers may have a range of interests which they are seeking to pursue through their participation in a CCP for which they may seek to engage the project scientists and software engineers, for which the relationships presented in this second category of literature may form a useful basis for analysis.

Chapter 3: Methodology

This Chapter describes how I set about answering my Research Questions. I chose to adopt a qualitative, ethnographic approach to studying CCPs, and this is justified in Section 3.1. There, I provide a brief overview of how ethnographic approaches have previously been applied by scholars to the study of science and scientists (subsection 3.1.1), and to the study of online communities (so-called *'virtual ethnography'* (Hine 2000), subsection 3.1.2): thus, it is a promising approach for studying the production of scientific knowledge by communities mediated primarily through the Internet. Researchers have found, however, that conducting virtual ethnography is not a trivial task of applying offline methods to online phenomena, and I discuss some of the issues that might arise in such cases (also subsection 3.1.2). These issues are then picked up in subsequent Sections in this Chapter where I discuss how I proceeded with my research.

Having set out my underlying approach, I then present my data collection in the subsequent four subsections. Each present a different source from which I drew my data: in each, I explain how I went about collecting these data, again considering how this was affected by much of it being about online phenomena and (in many cases) itself being collected by my use of online methods. I also discuss how I took into account ethical issues when collecting and storing data from each of these sources. I then set out how I went about analysing my data, taking into account the contexts in which they were produced and the impact upon their interpretation. The concluding Section of this Chapter looks forward to the rest of this thesis, explaining how my data will be used in my empirical Chapters, where I will present my findings.

Although this Chapter divides my research project into apparently consecutive stages, in reality I was engaged in all stages presented below at all times, with my approach to each stage shaping, and being shaped by, the other stages. I was continually reflecting upon and refining my approach in response to encountering unexpected barriers or previously unanticipated opportunities to take my research in new directions, or coming across phenomena of whose existence I had not hitherto been aware, or simply as a result of deepening of my understanding of the communities I was studying.

This need for such reflexivity comes from the very nature of social scientific research, as the researcher does not know in advance what they will find (Hammersley & Atkinson 1983). What this meant in practice was that rather than my data collection being completely separated from my data analysis, I instead analysed data as I progressed and used my findings to guide subsequent data collection. I reflected upon my experiences in collecting and analysing my data in order to refine and change the boundaries of my fieldsites, considering how far to trace the people, objects, and institutions I encountered in my research as they moved across various online and offline domains. Similarly, my ethical commitments were also subject to change over time, being shaped by (and in turn, shaping) my experiences in the field.

3.1 My approach to research: Online ethnography

In order to answer my Research Questions, I was interested in the practices that various individuals involved with the projects used to construct their relationships with others, how they negotiated their positions in the social hierarchies, and thus how these hierarchies were constituted. How did norms of behaviour and interactions emerge? How were they contested? How did they become stabilised? How were they reproduced across different situations? For instance, how did certain individuals or groups seek to impose such norms on others, whether and how was this resisted, and to what extent was this resistance successful? What communication practices were used, and in what contexts? What patterns of discourse emerged (for instance, ways in which things (objects, people, places) are constituted and classified by members of a project, or what tropes are commonly used)? What stories were told by the members of a project about the life of the project or the community that formed around it? What patterns of behaviour were particular to the individual projects I studied and why, and what were common across projects?

In other words, I was interested in understanding the cultures of CCPs. Ethnography is a qualitative approach that seeks to develop a bottom-up understanding of the cultural practices of the communities under study, and thus it made sense for me to adopt such an approach (Hammersley & Atkinson 1983). Its methods consist of: observing members of the culture being studied as they perform their usual activities and interactions with other members of the culture; interviewing these people; and collecting cultural artefacts for analysis. The aim is to provide what Geertz (1973) referred to as a *'thick description'* of the phenomena (rituals, practices, norms, values, discourses) that constitute the culture, in the sense of a detailed account that does justice to the emergence, stabilisation, reproduction, and inter-connectedness of these (p. 3).

Ethnography initially emerged from the field of cultural and social anthropology. In the early days, a researcher would focus on a single, remote, geographically-bounded culture alien to the researcher's own, and aimed to gain a holistic understanding of the culture that in which the entire lives of its members was embedded (for instance, Malinowski (1922) and Mead (1928)).

Since these studies, the techniques of ethnography have been taken and modified by researchers working in different domains. For instance, rather than seeing a single culture as completely bounded, more recent studies have stressed that all cultures are influenced by global forces, even those that might be considered remote (Burawoy 2000). Ethnography has also been used to study communities which are closer-to-home, for instance working-class communities in London (Hobbs 1998), marginalised communities in the western world (Wacquant 2008), and consumption practices in capitalist economies (Lien 1987). It has also been used to study communities of people whose members are also members of other communities (i.e. where the culture under study forms only part of its members lives), for instance, studies of particular organizations such as the IMF (Harper 1998) and studies of the workplace (Orr 1996). One particular sub-genre of workplace studies has been the study of scientists (both within, and outside, the laboratory).

3.1.1 Ethnography and the study of science

Citizen Cyberscience Projects are sites both where scientific knowledge is made and also sites where scientists and the public interact. Ethnographic approaches have been successfully employed within the field of Science & Technology Studies, both to study scientist-laypeople interactions and also to study science-in-the-making (see Chapter 2, Sections 2.1 and 2.2). These two groups of studies have shown how scientific norms and practices can be understood as cultural norms and practices, with groups of scientists (whether a research group in a laboratory, a department in a University, a discipline, or some other institution) as cultures bounded together by these practices. They have charted the emergence of these practices, how they are transmitted to and learnt by new members of the culture, how they come into conflict with norms and practices of other cultures (for instance, those of other scientific disciplines, or those of groups of lay people). They have also demonstrated how scientific knowledge claims in a particular discipline are constructed, contested, certified or rejected, and finally disseminated across that discipline (Hess 2001; Martin 1998).

Studies of science-in-the-making have employed a variety of ethnographic approaches. Those employing ethnomethodological approaches have documented how groups of scientists working together in a laboratory employ mundane practices to accomplish their day-to-day work (for instance, Garfinkel *et al*. (1981), and Lynch (1985)). Such approaches focus solely on the everyday practices of scientists within their laboratories without reference to any external factors.

Some other laboratory studies have also shown how scientists' orient their practices to the broader scientific networks in which they are embedded: their attempts to make sense of often messy and confusing events and observations are driven by the goal of greater recognition by their peers and, thus, acquisition of further funding to continue their work (e.g. Latour & Woolgar 1979). The impact on the very heart of scientific practice by factors external to scientists' laboratories has also been considered by studies which have brought ethnographic methods to bear on historical studies of science. For instance, Galison (1997) has situated the cultural practices of theoretical physicists within the historical development of that discipline, showing how changes in practices over time have been driven by changes in broader society.

Other studies, reflecting broader recent trends in ethnography, have engaged in multi-site ethnography, rather than restricting themselves to a single laboratory or discipline (for instance, Heath *et al.* 1999). This approach perhaps best exemplified by Knorr-Cetina (1999), who presented a comparative account of cultural practices of two groups of scientists in different disciplines, high-energy physics and molecular biology, finding that comparison of both sharpened her understandings of each:

> *'One may look at one science through the lens of the other. This "visibilizes" the invisible; each pattern detailed in one science serves as a sensor for identifying and mapping … patterns in the other'* (p. 4)

I set about developing my approach to research. Following Knorr-Cetina, I chose to focus on two research sites (i.e. two CCPs) which were sufficiently different to enable contrasts to be made but also sufficiently similar that meaningful comparisons could be made between them.

I was also eager to look at the broader contexts in which these projects existed, in order to understand how the practices that constituted the very basis of these projects were influenced, even driven, by external, global factors. I focussed on two contexts in particular. The first was the historical and social context of the scientific discipline in which a project is situated, in order to better understand how the disciplines emerged in relation to the concerns of society-at-large, with a view to understanding how the scientists' everyday practices in their projects are oriented towards these contexts.

The second context was that of the relationship between scientists and laypeople in society-at-large and I believed it was important to understand this. This was because both groups of people were likely to have come to the project laden with preconceptions about the other and about power inequalities between them. For instance, a layperson will have encountered science and scientific knowledge in numerous ways, suffused through every level of their lives, including newspaper reports about scientific research, technoscientific disasters (e.g. the Chernobyl disaster), educational background (school, university etc.), popular science texts, medical treatment, watching the weather forecast, and using technologies on an everyday basis. Thus, before getting involved in a project, a volunteer is likely to have an already highly-developed view of science and scientists relating to, for instance, trust, the status of scientific knowledge relative to other forms of knowledge and the value of such knowledge to society and themselves. Similarly, scientists are likely to have extensive prior experience of engaging with laypeople, be it doing formal public engagement activities, or simply explaining their work to an interested friend or family member.

Thus, ethnographic research of two CCPs, accompanied by situating these projects in their broader social and historical contexts, seemed a promising approach to understanding how relationships between volunteers and scientists, and amongst volunteers themselves, are constituted within such projects. However, unlike the ethnographic accounts of scientists that were discussed above, these projects are mediated to a large extent by the Internet, which raises complicating factors about how to proceed with the ethnography.

3.1.2 Online ethnography

Since the early 1990s, scholars have started to apply ethnographic methods to the study of communities which are largely or exclusively mediated by the Internet, recognising that these communities develop cultural norms of behaviour and communication (Escobar 1994). Early ethnographic studies of online communities include a study of social relations within a community of soap opera fans mediated by a UseNet mailing list (Baym 1995) and the negotiation of online identities and the formation of community on a bulletin board intended for lesbians (Correll 1995). These researchers, along with others who have since conducted *virtual ethnography* (Hine 2000), found that transferring ethnographic approaches from offline contexts to online communities is not a trivial matter; instead, the particular constraints and affordances of mediated communication compared to face-to-face communication has a profound impact on online ethnographic research. Of particular importance to my research is the impact along the following two dimensions (Markham 2011):

1) **The very nature of the communities being studied,** in the sense of the norms of behaviour and the performance of identity and community by those involved in online communities;
2) **My own research methods,** including ethics of research, the types of data that I can collect, what these data represent, and how to handle and analyse these data.

The aim of this subsection is not to lay down a prescriptive set of principles or methods for online ethnography – given that its inception is still relatively recent, there is still a great deal of uncertainty and debate about how researchers should proceed (Markham & Baym 2009). Instead, I present a brief summary of issues that have been encountered by online researchers with respect to the two dimensions mentioned above and that informed my sensibilities as I proceeded with each stage of my research.

Two particular issues to highlight here are: 1) there can be substantial differences between the nature of face-to-face communication and that which is mediated by Information Communication Technologies (ICTs); and

2) the Internet might be understood to fold physical space (in the sense of enabling close linkages between individuals who are geographically distant), compared with face-to-face communication.

Face-to-face communication has a number of features which may be missing from forms of ICT-mediated communication[9], and these can have a profound impact upon the ways in which individuals seek to make themselves understood and how they interpret what others are saying, the ways in which they construct and enact identity and community in online settings, and the relationship between an individual's online and offline lives. Many such issues will also impact on a researcher's study of online communities, in the sense that much of a researcher's interactions with their subjects will be mediated: CCPs typically involve both scientists and volunteers who are geographically highly-distributed, meaning that, for instance, it is not feasible to interview them face-to-face and instead contact must be made using mediated methods of communication.

There are a number of possible consequences of these features of mediated communication, for both the nature of the online communities themselves, and for the researcher. One is that an absence of visual or audio cues means there are there are liable to be more misunderstandings and breakdowns in communication during mediated communication than in face-to-face communication (Olson & Olson 2000). In turn, a lack of such cues may lead to the development of new methods and norms of communication. For instance, *emoticons* are often used to convey emotions in text-based communication (Rezabek & Cochenour 1998).

The features of mediated communication may also have implications for the construction and enactment of identity and community online, and this in turn problematizes the relationship and flow between the online and offline in the sense that there may not be a seamless connection. In the case of identity, an absence of visual and audio cues can mean that, online, an individual may be cut free of offline constraints, affording a great deal of plasticity to their disembodied online identity (Turkle 1995). This can give an individual much greater control over how they construct their identity online compared with offline (Markham 1998). Online, they may be able to transcend the preconceptions they encounter offline regarding their gender, race, age, educational background, class, (dis)abilities etc. (for instance, see Haraway 2004).

This has a number of potential implications for these online personae and the communities they form, and also (from the point of view of the researcher) for studying them. For instance, it might be thought that because mediated communication often affords fewer channels for the conveyance of information, then online relationships

[9] Clark & Brennan (1991) have identified several key features of face-to-face communication which can be missing from, and impede, ICT-mediated communication. These include: the ability of participants to see and hear each other, which allows for additional channels of communication, for instance tone of voice and body language; all participants being in the same location which can aid communication, for instance through the use of physical objects as shared points of reference; and the fact that all parties are involved in communication at the same time, so there are no delays in the exchange of information.

are impoverished and lacking compared with those conducted offline (for instance, see Jones 1994). However, substantial research exists to suggest that, within the context of online communities, a person's online identity can, in fact, be highly meaningful to them even when this identity appears to bear little resemblance to their offline identity (for instance, see Dibbell 1993; Turkle 1995). This implies that online communities can be seen as sites of rich interactions between highly-developed identities. For the researcher, it also mandates them to take ethical considerations seriously as they pursue their study.

Another implication is consideration of the relationship between offline and online contexts. What is the role they play with respect to the other (Miller & Slater 2000)? For instance, how do online communities shape offline lives (e.g. in the case of CCPs, many volunteers end up meeting in real life, or find involvement with a project can inspire them to make life-changing decisions, such as choosing to study science at university)? From the perspective of the researcher, this means they should take into account how to trace volunteers as they move across contexts, suggesting that they might seek to find out the broader social contexts (both on- and offline) in which they are embedded, and how they understand the importance of their involvement in the projects to their lives in general (for instance, Haythornthwaite & Wellman 2002).

In summary, then, the features of communication afforded by ICTs mean that online communities may potentially be very different in nature to those based more on face-to-face communication, in terms of how those involved are able to construct their identities, to negotiate social hierarchies, to communicate and to make themselves understood by others. These features also have implications for how a researcher will be able to study and interpret such communities. However, the issues discussed above are not the only ones that must be taken into consideration when studying ICT-mediated communities.

Another affordance of mediated communication is that it allows for the folding of physical space compared with face-to-face communication, in the sense that online communities can be formed amongst geographically-dispersed individuals and also that an individual may find it quite easy to become part of more than one online community simultaneously. Similarly, the Internet folds space for the researcher as well, allowing them to make contacts across the globe and to follow subjects across cyberspace.

Two implications of the Internet's impact upon spatiality will be considered here. The first is that individuals can be understood as being part of online networks of linked communities, rather than being geographically confined to one. For instance, in the case of climateprediction.net, a volunteer can easily sign up to another BOINC-based project and post in its forums (thereby joining another community) or join a team and post on the team's forums. Thus, a volunteer's online identity is possibly being shaped not only by (and, in turn, helps to shape) the climateprediction.net community, but also through their interactions and relationships in other communities. As a result, I chose to follow individuals and connections across online communities, rather than following the more traditional ethnographic approach of drawing clear lines around sites of research (following Hine 2007). Of course, not only does the Internet provide greater affordances for individuals to become part of many different communities, but it also provides greater affordances for the researcher in turn to follow the links between these communities – for instance, the forums of other BOINC-based projects are just a few mouse clicks away from the climateprediction.net forum.

The second impact of the folding of geographic space is that it may lead to the formation of communities of individuals with backgrounds far more diverse than in offline communities. People can be drawn from many multicultural and international backgrounds, which can give rise to issues which may lead to breakdowns in understandings amongst community members (Hinds & Weisband 2003). Such issues also need to be borne in mind by researchers as they, too, may be studying individuals from different cultures to their own.

One such issue is that in the cases of CCPs, volunteers from different countries may have very different understandings about the role of science in society (for instance, Jasanoff (2007) found differences in public cultures of science between countries) and thus may struggle to converge on shared understandings about the purposes of the project, leading to disagreements. They may also have very different expectations regarding etiquette, for instance what might seem acceptable behaviour in the culture of one volunteer may be regarded as rude or offensive within the context of another's (Hinds & Weisband 2003). Finally, an international community is likely to have more members who are communicating in a second language, which can promote misunderstandings or inhibitions in communication (de Rooij *et al*. 2007).

In subsections 3.1.1 and 3.1.2, we have seen that ethnography is a promising approach for the study of communities that form around CCPs. How I proceeded in practice is set out in the next Section.

3.2 Proceeding with field work

In Chapter 1, Section 1.3, I explained why climateprediction.net and Galaxy Zoo were chosen as case studies. I undertook to gain a thorough understanding of the broader contexts (historical, institutional, disciplinary, social and political) in which both projects are embedded, as well as engaging in a programme of collecting and analysing data about the projects themselves (primarily interview data with project personnel and observation of the projects' online forums). In the case of each project, I would spend some time in the field furthering my understanding of the project and collecting data, and then some time away analysing the data, reflecting on my progress and then planning my next steps, before re-entering the field: thus, the various stages set out in this Section can be seen to be co-shaping each other rather than independent or discrete. Underpinning this research was a sense of ethical responsibility to those whose communities I was studying, and it is to this that I turn first.

3.2.1 A note on research ethics

In social research, ethics has been an important component of designing, carrying out and writing-up fieldwork projects for many decades. It has been widely-recognised that studying the cultures and societies in which individuals are embedded has the potential to cause significant emotional harm to these individuals and societies. For example, the researcher may use methods that those studied find intrusive, or may produce accounts of their fieldwork that result in substantial societal disruption or exposure of individuals. As a result, there has been much of consideration about how social researchers should proceed, drawing on a variety of ethical theories and how they might apply in different situations (for instance, Ess 2006).

Translating these approaches into an ethical framework was not a straightforward task. I chose not to develop a rigid framework before entering the field, for two main reasons. The first is that, given the relative newness of social research on the Internet and the difficulties inherent in transferring offline research methods to online contexts, there is a marked lack of consensus regarding how to develop such a framework (Buchanan 2011). The second is that, given my methodological commitment to try and understand the communities I was observing in their own terms, I did not want to pre-judge what would be appropriate behaviour as a researcher. Instead, following Markham's (2006) call for a *'reflexive'* approach to online social research (p. 37), I entered the field with a set of sensibilities that I would reflect upon and refine, and how these were applied to my research will be discussed in subsequent subsections.

These sensibilities were informed by existing literature on the ethics of internet research (a particular influence here was the report of the Ethics Committee of the Association of Internet Researchers, which is widely-cited (Ess & AoIR 2002)). I was especially sensitive to the notion that these online communities were potentially very meaningful for the participants (both professional scientists and software engineers and volunteers). From the perspective of the scientists and software engineers, most had a high professional stake in their project, and I was therefore concerned that my research would not prove disruptive to this.

From the perspective of the volunteers, I became aware of the importance of the online community to participants in terms of socialising and online relationships and sometimes the opportunity to construct identities in ways that were not necessarily available off-line (for instance, one disabled participant explained to me that nobody in the forum knew they were disabled and they valued not being treated differently from others as a result of this). It also became apparent that the online interactions were important to subjects in their offline lives as well. For instance, one volunteer on Galaxy Zoo organised a spinoff research project using threads on the project's forums, and has now used this as a basis for starting a PhD at the Open University, whilst another has been encouraged to pursue science communication as a career, encouraged by responses to their forum posts discussing the project science. These ethical concerns had a number of implications for how I approached my research at all stages. These will be discussed where appropriate in the following subsections.

3.2.2 Interviews

Interviews comprised a very important part of my data. During the course of my research, I conducted interviews with 22 scientists and software engineers and 36 volunteers across the two projects. These interviews were conducted using a variety of media: face-to-face, Skype, Instant Messaging, and email. All interviews were semi-structured (Arksey & Knight 1999). What this means is that a schedule of questions was drawn up in advance of the interviews, however I was able to deviate from that list during an interview, for instance by asking questions out of order, and following up interesting and unanticipated lines of inquiry that arose. Such an approach combines the benefits of structured interviews where a list of questions is strictly adhered to (and thus allows for comparisons to be made across interviews) with unstructured interviews where there is no schedule of questions (which gives the researcher the flexibility to follow-up unanticipated lines of inquiry).

My ethical concerns guided how I proceeded with recruiting and conducting the interviews. I set out to secure informed consent from those I interviewed to allow me to make a digital audio recording of the interview (where it was face-to-face or conducted via Skype) or to keep a transcript of the interview (when conducted using Instant Messaging or as a series of emails), and to use quotes from the interviews[10]. The privacy of interview subjects was protected by keeping all of my data (both digital recordings and transcriptions) in a password-protected area of my computer, and storing any printed transcriptions in a locked filing cabinet. Only I have seen these transcriptions, or listened to the interview recordings.

[10] See Thorne (1980) for some of the issues involved in gaining consent. Varnhagen *et al*. (2005) discuss particular difficulties in making sure that informed consent truly is informed in the context of online research, for instance that the fewer visual and audio cues in mediated communication (compared with face-to-face communication) can complicate the researcher's task of ensuring that understandings between researcher and interviewee are reached. I made an effort to take these concerns into account.

I recruited interview subjects in different ways according to whether they were volunteers or project scientists. I sent emails to every member of the team that I could identify (past and present), as well as to scientists who were not part of a project's core team but who had collaborated with members of the project on work relating to the project. I sent individual emails to each, in which I introduced myself, briefly introduced my research, and attached an Information Sheet giving more details about my research such as its motivations, methods and expected. I asked whether they would consider being interviewed by me for my project. If they replied in the positive, I then sent further emails to establish a time and date for the interview, and I asked them explicitly whether I could record the interview and use (anonymized) quotations and also whether they had any further questions about my research. I sought to conduct as many interviews face-to-face as possible, although this was not feasible for many who were located abroad. In the event, I conducted 22 interviews with Galaxy Zoo and climateprediction.net scientists, either face-to-face or via Skype. All interviews lasted between 30 minutes and 90 minutes, with the majority lasting 45-60 minutes.

In the case of volunteers, I faced one challenge in particular with recruitment, namely that those running the project were not comfortable with me contacting volunteers directly to ask whether they would take part in an interview. In order to avoid spamming volunteers with unsolicited requests, I instead devised a short online survey to recruit potential interviewees, which I promoted on the project forums with the support of forum moderators. Two drawbacks of recruiting using this method are, firstly, that it does not guarantee a representative sample of volunteers as participants are self-selecting, and secondly that it restricted interviewees to those who actively consulted the forum. This excluded those who, for whatever reason, chose not to participate in the forums or who no longer participate in the project. Given the circumstances, however, this form of recruitment was the best that could be achieved. Furthermore, given that I had undertaken to study the community that formed around a project, it did not seem to be too much of a problem not to speak with those who did not actively constitute this community.

This survey asked eight brief questions about the volunteer's involvement with the project, and also consisted of a box at the bottom where the volunteer could indicate if they wished to participate in an interview. As with my interviews with the project scientists, I sent an information sheet and an explanation of how interview data would be stored and used and a request for consent.

3.2.3 Observing the online forums

The second major part of my fieldwork focussed on observing the online forums on both project's websites. These forums are the main site of interaction amongst project participants. They are the ideal place to study the cultures of the communities that form around a CCP as participants construct their online identities, and norms of behaviour and social structures emerge, get contested, and are stabilised through these interactions. I consistently visited the forums throughout the course of my fieldwork. Initially, I simply read threads at random in order to try and get a feel for the overall structure of the forums and to begin to make sense of them. At first, it was very difficult to find my way around them and to get a handle on what was going on: in the cases of both projects, the forums are vast (as of 31 August 2011, there were approximately 540 000 posts in around 17 500 threads across 28 sub-forums on the Galaxy Zoo forums, whilst the climateprediction.net forums contained over 100 000 posts in nearly 13 000 threads across 36 subforums).

After getting my bearings, I undertook a more systematic approach to the forums. I attempted to read all threads to identify some for further analysis. In most cases, this involved skim-reading very short threads. I read those threads that seemed interesting more carefully, and I saved around 500 threads for further study and analysis. As I sharpened my understandings of what was going on in the forums, and I started to identify themes or issues common across a range of threads, I was able to search the whole forum using key words or phrases to find out what had been posted on a particular topic during the forum's lifetime.

As I progressed, I also identified certain individuals whose postings were particularly interesting. Each had a profile page which could be accessed by clicking on their name in the forum, and through this, I was able to view other posts in order to gain an appreciation of the nature of their interactions with other volunteers and to follow their activity across the project's forums. Furthermore, in the case of climateprediction.net, this profile showed other BOINC-based projects in which they were involved, and the team of which they were a member, and through this I could trace their activity across a number of projects and to observe how they interacted with volunteers on other projects' and teams' forums. This enabled me to gain a greater understanding of broader contexts in which they operated in the world of Citizen Cyberscience.

Aware of the potential impact of my research on the ways in which forum members might understand themselves and their relationships with others on the forums, and that they might feel uncomfortable with being analysed, I used the Personal Message facility to contact those involved in any thread I wished to quote in my thesis to ask for their explicit consent to be quoted and for the threads they were involved in to be studied, and I stressed that I would use pseudonyms, rather than their online names, to identify them.

3.2.4 Other sources of data

My two main sources of data were the interviews and the observations of online forums, but these were supplemented by a wide range of other data collected for analysis. These include pages on the projects' websites (information regarding the science underpinning the project, the history and development of the projects, details about the scientists and software engineers involved in the project, tutorials regarding how to take part in the project and, in the case of Galaxy Zoo, a blog covering a variety of subjects, including the latest news about new developments in the project). Other sources of data were publications by members of the project (including journal articles presenting scientific results based on volunteer contributions, articles recounting the experiences of project scientists in setting up and running projects, funding proposals, and the personal websites, blogs, and Twitter accounts of project scientists, and of two Galaxy Zoo volunteers).

3.2.5 Supplementary work

Of equal importance to the structured fieldwork exercises described above was a wide range of other activities that I undertook to immerse myself more in the worlds of the projects I was studying. The aim of these was to gain an understanding of the broad historical, political, social, and scientific contexts in which each project is situated. For instance, I read widely about the history of both astronomy and climate science, about the history of the amateur in science, and processes of professionalization and mathematization in science during the course of the early twentieth century. This helped me to relate the development and goals of the projects I was studying to those of the scientific disciplines and institutions in which they are located, and to the establishment and maintenance of cultural authority of science in general as society's primary source of knowledge. Many of the texts I read are cited in the Chapters in which I present my findings.

I also sought to understand the science underpinning the two projects, and studied undergraduate textbooks and reading lists relating to astronomy and climate modelling. As I have an undergraduate degree in mathematics specializing in statistics, I found it interesting to learn about the mathematical models and statistical techniques involved in the science. This gave me a better appreciation of the projects' work, and meant I could follow scientific papers and conference presentations that resulted from volunteer contributions. I also attended a few conferences in both disciplines, which gave me an overview of current research trends, and thus how the work of the projects is oriented to the cutting edge in the rest of their fields.

Finally, in order to further my understanding of how the projects fit into contemporary societal and political currents, I followed coverage of both astronomy and climate science in newspapers, on current affairs programmes, and in popular science publications. For example, during my fieldwork on climateprediction.net, the *Climategate* scandal broke, where private emails of scientists at a leading climate change research centre were published which suggested that some results had been suppressed to avoid undermining the theory of man-made climate change (Nerlich 2010). It was fascinating to follow the debate across the media and how the scandal was marshalled by various social actors to support their own agendas. Reaction to the event helped to reveal the contours of public support or scepticism regarding climate science, which helped me to understand the challenges that climateprediction.net scientists face in recruiting and retaining volunteers.

Studying the contexts in which the projects are situated complemented my more structured data collection, and proved particularly useful when I turned my attention to interpreting and analysing my data.

3.2.6 Data analysis and writing-up

I attempted to analyse data and write up some of my findings as I went along, which allowed me to explore preliminary ideas about the communities I was studying, to present these in various public settings (seminars, conferences, and a journal article): this allowed me to incorporate feedback into my on-going fieldwork.

Ethnography involves seeking to understand the norms and practices of the culture under study in the terms of the members of that culture, rather than imposing upon it the researcher's preconceived notions. As a result, I adopted a bottom-up approach, allowing theoretical concepts and categories to emerge from the data. The standard approach for this in ethnography is *grounded theory* (Glaser & Strauss 1967). The first stage was to read the data closely, with the aim of identifying patterns of discourse across sources including how objects and people are grouped together and how these groups are placed in relation to each other, how concepts were constructed, and how themes were developed (Potter & Wetherell 1994). My data were then coded according to these themes, categories and concepts. I refined this framework through an iterative process: I would code my data according to the framework in order to see how well it matched the ways in which those I was studying made sense of the world and each other, which in turn would lead to adjustments of my framework (Hammersley & Atkinson 1983).

Additionally, I made a number of attempts to develop classifications of participants, using a variety of information (e.g. that available on the participants' profile pages on the projects' websites (for instance, in the case of climateprediction.net, the volunteers' credit scores, team membership, and the other BOINC-based projects in which they are involved), and the types of forum threads in which they were active). The idea behind this was to attempt to match different types of discourse (i.e. different ways of constituting the world) to different classes in order to gain a greater appreciation of how social structures and hierarchies emerged, are contested, and sustained in the communities of the projects under study. Again, these groupings were refined iteratively to produce a better fit with the modes of discourse employed and in conjunction with the framework described above.

At all points in the analysis, it was important to keep in mind that the content of each piece of data does not reflect an underlying reality; instead a piece of data is a contextually-produced artefact of sites of social interaction (Abell & Myers 2008). Such a context is dependent on a number of factors. One factor that is particularly important when considering the context in which a text was produced, namely that the author produces their text in anticipation of the effect that it will have on its audience, i.e. it is produced in anticipation of how its audience will respond to it, with the author seeking to influence and persuade towards a particular end (Bakhtin 1981). As a result, it should not be analysed with the intention of trying to reveal a true underlying reality, but rather in the context of what the author was hoping to achieve.

Thus, I was very careful when handling a piece of data, engaging in detailed consideration of the context in which it was produced. That a piece of data is contingent meant that I needed to engage in triangulation. This means that, rather than relying on a single piece of data when studying a particular phenomenon, I should instead cross-check (*'triangulate'*) it with data relating to the phenomenon that comes from a variety of other sources (O'Donoghue & Punch 2003). The idea is to ensure that conclusions drawn about a particular phenomenon from data are not biased by the context in which the data are produced.

From an ethical perspective, I needed to be very sensitive to the possible impact of the way in which I presented my findings. By studying communities so closely, there is always the possibility of the researcher exposing exploitative power relations, abusive or devious practices, subterfuge and so on. Even worse, there is the possibility that the researcher suggests the existence of such things where they do not exist. One theoretical repertoire used to interpret data might suggest one explanation for what has been observed, whilst another theoretical approach might suggest another explanation altogether. Thus, when analysing and interpreting my data, I was conscious of how I might represent various people in a way that was fair, accurate and non-disruptive.

3.3 Conclusion

In this Chapter, I have set out how I proceeded with my fieldwork, describing underlying considerations (for instance, my choice of ethnography and how to apply such an approach to the study of science-in-the-making in online sites of interaction between scientists and members of the public, and how I might proceed ethically) and how these informed (and, indeed, were shaped by my experiences of) my everyday research practices (data collection and analysis). Very frequently, my expectations of what I would find were confounded, causing me to reflect upon, and adjust, how I was conducting my research.

As a result of my studies, I learnt a great deal about the communities forming around two CCPs, as well as the broader contexts in which these projects and their communities are embedded, and my findings are presented in the three Chapters subsequent to this.

Chapter 4: The Galaxy Zoo project: Challenges past and present

4.1 Introduction

The following two Chapters present findings from my Galaxy Zoo case study. In particular, they are concerned with the relationships between scientists and volunteers in the Galaxy Zoo project and the way in which these relationships are constituted. By studying the discourse of the project (or, more strictly, the members of the project's core team), it appears that there is an ethos that underpins the norms in the project that govern how volunteers are to be treated and the core team members' interactions with the volunteers, and guide decision-making processes within the project. Henceforth, I shall refer to this as the *Galaxy Zoo ethos.* This ethos is mobilised by core team members, either explicitly or implicitly, to explain their actions and to justify developments in the project (such as new features being added, or changes in policies regarding the running of the project). The Galaxy Zoo ethos appears to be founded on the assumptions that the contributions of volunteers should be regarded as gifts to science which involve sacrifice (time, effort etc.) on the part of the giver, that the scientists have a moral imperative to acknowledge and respect this, and that they value all volunteers equally.

My core argument is that this ethos did not, however, exist prior to the decisions that the core team had to make during the setting up and running of Galaxy Zoo. Rather, I will attempt to show how it has emerged during the course of the project, how it has been shaped by (and in turn shapes) the interests of various project members, and, furthermore, how the technical aspects of the project have driven (and been impacted by) this ethos and these interests. At various points during the course of the project, those involved have had to make a series of choices (about the scientific work of the project, whether to become and remain involved in the project, policies governing the running of the project etc.) amongst alternatives available to them at the time. They have evaluated these alternatives in terms of considering how they might contribute to the pursuit of the actors' interests at the time, to what extent they accord with the ethos as it was at the time, and to what extent these alternatives were afforded by the material and social configurations that existed at the time (i.e. the ease with which different alternatives could be pursued). By choosing one alternative ahead of others, the actors involved would reaffirm some norms of behaviour whilst marginalizing others. This process would also open up further alternatives that might be pursued in the future, whilst closing down others, creating new configurations of choices to be made by social actors. Through this process, the Galaxy Zoo ethos emerged from amongst a range of possible alternatives and became stabilised.

In order to clarify what is meant by this, I will now focus on theoretical approaches to the study of the relationship between the technological and the social and discuss how these have informed my analyses.

4.1.1 Technology and society

This Chapter (and subsequent Chapters) considers the CCPs as sociotechnical systems, in the sense that these projects are both social (there is a community that forms around, and is mediated through, the projects) and are technologies for the doing of science. Furthermore, these two aspects are related, and indeed, might be seen as inseparable (because the work of a project would be impossible without a community to design, build and run it and process the project data, and conversely, the formation of the community depends on the existence of the project and the means it provides for different people to engage with each other).

Technological developments have often accompanied social changes, raising the question of how these might be linked. A large number of analyses of technological and social change have accounted for the latter (social) in terms of the former (technological), seeing technological change as the driving force behind the reconstitution of social relations. This has been labelled *technological determinism*, and can be seen in many accounts, both historical and social. Such an approach might, for instance, see changes in gender roles during the twentieth century as being driven by the development of so-called labour-saving devices like the washing machine leading to the inevitable liberation of women from endless domestic chores and thus freeing them up to pursue careers, and so on (Coen-Pirani *et al*. 2008). Technological determinism can also be seen in many accounts of the social and economic impact of the Internet. For instance, Benckler (2006) argues that the widespread availability (and relative low cost) of computational technologies and Internet access is taking power away from traditional institutions and businesses in the knowledge and creative industries and distributing it amongst a wide number of people, including (particularly) amateurs. Amongst the examples cited by Benckler are crowdsourcing in the production of Wikipedia, and potential volunteers for CCPs choosing which projects to support and thus what scientific work is done.

During the 1980s, however, some scholars of technology became dissatisfied with this approach, for instance, because it does not account for why some technologies achieve a wide uptake whilst other, apparently similar technologies, fail to do so. For instance, Deuten (2003) asks why the VHS format became a near-universal standard for videos whilst Betamax faded into obscurity. Inspired by social constructivist approaches to the study of scientific knowledge, these scholars developed an approach known as the *Social Construction of Technology* (SCOT) which sought to account for the development and uses of technology in terms of the social, arguing that a technology offered a limitless range of ways in which it might be used. Thus social actors could use and shape a technology in order to pursue their own social interests, rather than being mandated to use it in a particular way by it having certain essential features (Pinch & Bijker 1984).

SCOT was, however, seen as unsatisfactory by many: it seemed counter-intuitive to deny a role for material agency in accounts of the relationship between the technical and the social – after all, a gun would seem to offer somebody a more effective way of killing than a rose (Kling 1992). One attempt to chart a course between the two poles of the essentialism of technological determinism and the radical relativism of SCOT was Actor-Network Theory (ANT), which sought to achieve a symmetry between the material and the social by assigning interests to non-humans (for instance, other living organisms such as bacteria and scallops, pieces of technology, scientific instruments) and thereby turning them into social actors as well (Callon 1986a; Latour 1987)[11]. The underlying idea of ANT is that an actor is motivated to pursue their own interests, and thus attempts to build a network through enrolling other actors by convincing them (either through persuasion or coercion) that this enrolment will enable them to better pursue their own interests. For instance, Latour (1988) discusses how the French microbiologist Louis Pasteur assembled a vast network including various people, institutions, scientific instruments and methods, and microbes across France in order to demonstrate the microbe theory of disease and also to secure recognition as the originator of this theory.

Through ANT, the notion of what it means to say that an actor has interests became problematized: rather than an actor's interests being fixed, existing independently of social action, they are instead regarded as indeterminate and subject to change. Callon (1986b) described the process of *intéressement,* where an actor attempts to define and shape another's interests so that the latter would be more likely to see enrolment in the former's network as a desirable course of action. The former might, for instance, attempt to show the latter how they would not be able to successfully pursue their current interests (thus creating a space for new interests to be introduced) or to alert the latter to possibilities of which they had previously been unaware.

In other words, ANT suggested that the material and social were not fixed, and thus socio-technical systems are hybrids that emerge as a result of interactions between the social and the material. Furthermore, it suggested that the interests of actors are not fixed (and thus do not determine the way in which these systems emerged) but both constitute and are constituted by these systems. However, there was some dissatisfaction with ANT, and I will deal with two criticisms in particular which relate to my research. One is that it is monovalent, in that it sees a social actor's identity as singular across all possible contexts (Mol & Law 1994). This criticism of ANT is relevant to CCPs because they bring together different groups of people who may be situated in disparate contexts, and as I discussed in Literature Review, they are likely to be embedded in different epistemic cultures

[11] Another attempt to produce accounts of technology which privilege neither the social nor the material is that of Hutchby (2001) who advanced the concept of affordances, suggesting that a technology might enable an actor to take particular courses of action whilst con straining them from taking others, without determining which courses of action to take.

and thus liable to interpret the same thing in very different ways. In this sense, an ANT-based approach would offer limitations for studying the CCPs.

The second criticism of ANT is that it sees social actors as motivated only by self-interest and thus willing to manipulate others, and does not leave room for an actor's behaviour to be accounted for other than by this (Amsterdamska 1990). In this approach, the Galaxy Zoo ethos might be understood purely as a rhetorical device to conceal their manipulations of volunteers in pursuit of their own self-interest. That this would be an unsatisfactory approach for analysing my case studies was borne out, in the sense that I found that the project scientists have genuine respect for volunteers and do not simply view them as a means to their own ends.

The first criticism highlighted above has been addressed by Mol & Law (1994). Whilst an ANT approach would see a thing as having a single social identity constituted by the links it has with other things in its network, Law & Mol discussed how a thing's identity is highly context-dependent and can thus be seen as having a multiplicity of social identities across different contexts (see also Mol 2002). A thing's identity in a particular context is constituted by the relations that it makes in that context. For instance, they discuss the example of anaemia and how its social identity varies between the Netherlands and Africa. In the Netherlands, it is detected through laboratory tests, defined as a small variation from the norm, and is not seen as a particularly serious condition: this is because such testing methods, which can detect small variations, are readily available in the Netherlands. Furthermore, because it is easy for residents to follow a diet with sufficient iron in it to avoid anaemia, it is usually only present as a small deviation from the norm and, at such levels, has little impact on the sufferer. In Africa, by contrast, such laboratory techniques are often not available, anaemia is diagnosed by observing physical symptoms and, in addition, because it may be more difficult for a resident to have nearly enough iron in their diet, anaemia is more likely to be severe (and thus physically observable) and may have serious repercussions for the sufferer and for their community.

The second criticism of ANT highlighted above has been addressed by de Laet & Mol (2000). They have suggested that actor-networks are but one of many different forms of social topologies by which to account for the development of sociotechnical systems. For instance, in their study of the Zimbabwe bush pump, Mol and de Laet provide a description of how its spread and wide uptake is not due to the assemblage of a network by an individual or a group, but rather because it has proven highly-adaptable to the various local circumstances in which it is used (for instance, is highly durable, it can be easily-repaired, and its mode of construction brings together the whole community working together, in turn fostering a sense of collective responsibility for its maintenance). This way of accounting for the development of sociotechnical systems opens up spaces for individuals to act in ways other than is allowed by ANT. For instance, in the case of the bush pump, the designer and manufacturer is found to be a modest, self-effacing individual who does not seek success through social manipulation – far from assembling a network and attempting to define the interests and roles of, and assign actions to, each member of the network, he is sometimes surprised to discover how the bush pump is used and how it is maintained in practice.

Some points to note here: the first is that the relations in which anaemia is embedded (including diet, the social relations of the sufferer to others in the community, the technologies and expertise available for detection) in each context gives rise to very different social identities for the disease. Secondly, these identities can develop and change, and also become stabilised, over time: some relations are strengthened and become deeper as they are made and remade over time, whilst the making of other relations can in turn open up the possibility of making new relations with other things. For instance, in Africa, the absence of laboratory testing methods and the more severe presentation of symptoms leads to doctors and nurses developing their skills at being able to diagnose anaemia through visual observation; conversely, by becoming more skilled, they are able to make diagnoses more easily which both strengthens anaemia's identity as something that is diagnosed in this way, and also contributes to anaemia being recognised as widespread and a serious problem in society leading to a greater focus on it by policymakers, NGOs, charities etc.

This can also be extended to the study of the development of the ethos of a community, in the sense that this ethos can be seen as being constituted by the relations (with words, objects, people etc.) that are made more available and reinforced, or made less available, in a particular techno-scientific configuration. For instance, new computational and Internet technologies present challenges in terms of establishing new ethical frameworks regarding, for instance, notions of privacy as they open up new possibilities and risks that have simply not been anticipated (and thus addressed) before (see Chapter 3, subsection 3.2.1).

Applying these insights (and insights discussed further above) to my analyses of CCPs, I will bear a number of points in mind as I proceed through the Chapter. The first is that the interests of various social actors (the scientists and institutions involved in the project, the volunteers, the broader scientific community and other members of society and so on) are not fixed in advance, and do not exist independently of the sociotechnical systems in which the actors are embedded: instead, they are shaped by, and in turn shape, these systems. Similarly, as Carusi and De Grandis (forthcoming) argue, ethos does not exist prior to, and independently, of sociotechnical systems, nor of the interests of the various actors involved. Rather, there is a co-shaping of these elements. Finally, these elements emerge, change and are stabilised over time in a space of relations, and this involves actors making choices between the different configurations of relations that are available at a particular point in time. As new relations are made, certain new configurations of relations become possible whilst others might be closed down and, furthermore, particular relations are strengthened as they are made and remade. In other words, interests and ethos can be seen as socio-technically located, and this is the theoretical position I adopt in my analysis of CCPs in this thesis.

4.1.2 The structure of my argument

This argument is split across the following two Chapters as follows: in Chapter 4, I will consider the interests of the team involved in the setting-up and running of Galaxy Zoo, how they have pursued these interests to date through Galaxy Zoo, how these interests have changed over time as they have discovered (and pursued) previously-unanticipated opportunities afforded by the project, how this has impacted on the way the project has been developed, and the potential threats to these interests that these team members see looming on the horizon. In the subsequent Chapter, I will describe the features of the Galaxy Zoo ethos and attempt to show that, rather than being fixed and inevitable, they have emerged as a result of choices made from a range of alternatives which the core team believed would better enable themselves to pursue their emergent interests.

Of course, no group of social actors exist in a vacuum; instead, they are dependent on, collaborate with, and are in competition with, other actors. Thus, I need to contextualise Galaxy Zoo in order to show: 1) how the projects' goals became oriented towards the goals of cosmology; 2) how these goals relate to the discipline in which cosmology is situated (astronomy); and 3) how the study of astronomy has achieved legitimacy in society-at-large. This will be covered in Section 4.2, where I provide a brief historical survey of astronomy, how cosmology's aspirations for most of the 20^{th} century has been to raise its status relative to other astronomical subdisciplines, and how Galaxy Zoo has presented new opportunities to achieve this.

In response to these contexts, I then describe the development and work of the Galaxy Zoo project itself (subsections 4.3.1 and 4.3.2). I also present the *Citizen Science Alliance* (CSA), an umbrella organization of CCPs (collectively known as the *Zooniverse*) set up by key members of the Galaxy Zoo core team, and argue that these individuals can be understood as being interested in the possibility of leading a growing Citizen Cyberscience movement through the CSA. In particular, because the core team had not originally anticipated the scale of Galaxy Zoo, this possibility was not apparent to them when they initially launched the project, and this is an example of how their interests have evolved. Further, I argue that these individuals regard the continued success of the Galaxy Zoo project itself as critical to the potential success of this movement, and the pursuit of this new interest therefore, in turn, strengthens their interest in securing the project's future (subsection 4.3.3).

Drawing on my data from interviews with people involved in Galaxy Zoo and other sources of information, I discuss the challenges the core team sought to address when establishing Galaxy Zoo as a successful scientific project, and the challenges they expect to face in order to sustain this into the future. One such challenge is the core team's dependence on enrolling other astronomers to use the project data in their own research in order to justify the project's continued existence. By situating these astronomers in the broader context of cosmology, it can be seen how the core team have come to an understanding that their continued ability to enrol these astronomers relies almost exclusively on the generation of data of a scale and quality unmatched by other sources of data and not on the other factor that might be assumed to attract them to Galaxy Zoo, namely the public engagement aspect (subsection 4.4.1). Linked to this challenge is another, namely the project's struggle to gain scientific credibility in the face of scepticism from the discipline-at-large, and the hard work and numerous tasks undertaken by core team members to this end (4.4.2).

Having thus discussed how Galaxy Zoo became established, the successes-to-date which have resulted from it, and also set out the future aspirations of those involved in the project, I then turn to contemporary developments in cosmology (Section 4.5). In particular, I describe methods that are currently being developed and implemented to create data sets of a scale and quality that far outstrip those which can be produced by Galaxy Zoo, and are seen by core team members as threatening to supersede Galaxy Zoo in terms of contributing to cosmology's core disciplinary interests. In turn, they are worried that this will render their project obsolete, because they will not be able to continue enrolling astronomers to use the data generated by the project. Under such conditions, one might ask why the members of the core team are so concerned with continuing the project in the future – after all, many of them are cosmologists and thus might be assumed to want to move on and keep up with the rest of their field. I argue that this is because their interests have changed both over time and in the shifting network of relations in which they are embedded (see subsection 4.1.1), and should no longer be understood only in the context of astronomy, but also in terms of their aspirations with the CSA, and Galaxy Zoo's perceived importance in establishing this.

Finally, I summarize the key points of this Chapter, in particular how the core team have perceived (and addressed) the challenges to the pursuit of their interests-to-date and to the pursuit of their interests in the near future, This provides a foundation for the next Chapter, where I present an account of the development of the Galaxy Zoo ethos.

4.2 Galaxy Zoo in context

In this Section, I discuss the contexts into which Galaxy Zoo emerged, describing the major challenge that the project set out to address and explaining why this challenge was so critical for cosmology (the subdiscipline of astronomy concerned with the origin, evolution and the ultimate face of the universe and in which Galaxy Zoo is located) as a whole. I first present an overview of the history of astronomy, explaining how cosmology has had a lowlier status than other astronomical subdisciplines because statistical confirmation of most of its basic conjectures has been lacking and how cosmologists therefore have been concerned with raising the status of cosmology. This has driven the development of methods over the past fifty years to gather much greater quantities of cosmological data (in this case, images of galaxies) to provide a firmer empirical foundation for cosmological theory. I will explain how the two approaches that had been available for processing this data (professional astronomers classifying the images by hand, and automated image recognition software) have so far been inadequate for the task: thus, the challenge facing cosmologists was the question of how to handle this data deluge. This leads into the subsequent Section, where I will describe how a team of cosmologists at the University of Oxford set out to respond to this challenge through the development of Galaxy Zoo.

4.2.1 Cosmology and the history of astronomy

Cosmology is a subdiscipline of astronomy, a discipline that sets out to answer questions that have fascinated people for millennia. Public, financial and material support for the study of the universe has relied on a number of different factors over time. The cosmos was an object of study for the Ancient Greeks, and was strongly related to their religious practices, and transcendental reasons have usually underpinned subsequent public and institutional support for the study of the universe (Bowler & Morus 2005). For instance, in the Middle Ages, the Catholic Church supported the study of the planets and stars as holding the key to unlocking some of the mysteries of God’s creation (although the study of the solar system was repressed during and after the life of Nicholas Copernicus after he suggested that the Earth, and thus Man, was not at the centre of the universe). In more recent decades, despite a decline in belief in organized religions in western society, there has been a strong continuation of support of the study of astronomy as humans seek to understand their place in the universe (Smith *et al*. 2010). Astronomy has also received support for more political or utilitarian reasons, such as its usefulness to navigation (Portolano 2000), military applications (Devorkin 1980) or practical spin-offs from the space programme (Goldin 2004). It has also played an important role in nation-building, and in international relations: projects such as building observatories or the space race provide a great deal of symbolism and status for nations. For instance, Portolano (2000) describes how the building of an observatory in the early United States was presented as an opportunity for the USA to raise its status compared with European nations. In order to study these various questions of interest, a number of subdisciplines of astronomy developed over time, each studying different regions of the universe, or different objects in space. Along with cosmology, these sub-disciplines include: solar astronomy (the study of the Sun), planetary astronomy (planets in our Solar System), lunar astronomy (the Moon), and stellar astronomy (which studies the life cycle and properties of stars).

During much of the twentieth century, the central question cosmology sought to answer was whether the universe was steady-state (in the sense that it had no origin, no end-point and remained the same size throughout time) or did it originate with a Big Bang and, if so, will it continue to expand forever or will it begin to contract (and thus end in a Big Crunch)? In the closing decades of the twentieth century, the scientific consensus has come down firmly against the steady-state theory and in favour of the Big Bang. However, a number of major questions of interest remain unanswered, for instance relating to the underlying structure of the universe and the rate at which it is expanding and at what point it might begin to contract.

The astronomer Edwin Hubble, in the 1920s, established that the key to answering the questions about the universe lay in understanding the evolution and structure of galaxies (Hubble 1926). Hubble devised a classification scheme for galaxies, and this has underpinned work on galaxy morphology since. In this scheme, galaxies are classified according first to whether they were elliptical or spiral galaxies. In the case of elliptical galaxies, they were then further divided depending on whether they appeared circular or more elliptical in the sky. Spiral galaxies are classified according to how tightly their spiral arms are woven around their centre, and also whether (and how large) is a bulge at the centre of the galaxy. This classification scheme is known as *'Hubble's tuning fork'* (Figure 4.1).

4.2.2 The lowlier status of cosmology in astronomy

Answering these questions in the terms laid down by Hubble, however, has not been a straightforward task. The result of this is that compared with other sub-disciplines of astronomy, cosmology has often been regarded as having a somewhat lowlier status: for most of the twentieth century, cosmologists were not able to provide solid quantitative evidence to support its assertions and conjectures.

A major feature of the professionalization of science in general at the start of the twentieth century was a greater focus on mathematical methods (Porter 1996), and astronomy was no different, as it helped professional astronomers to differentiate themselves from the mass of amateurs. One reason is that quantitative methods helped to endow their science with a greater sense of objectivity. Another was that it was a form of discourse available only to those who had received mathematical training as part of a university education, and thus served to exclude amateurs from contributing (Lankford 1981). Cosmologists' struggles to deal in such terms mean that their field, in turn, struggled to attain the same status as other sub-disciplines of astronomy: for instance, one leading astronomer dismissed cosmology thus: *'there are only two and a half facts in cosmology'*, implying that the rest is merely speculative (quoted in Kragh (1996, p. 320). See also Bowler & Morus (2005)).

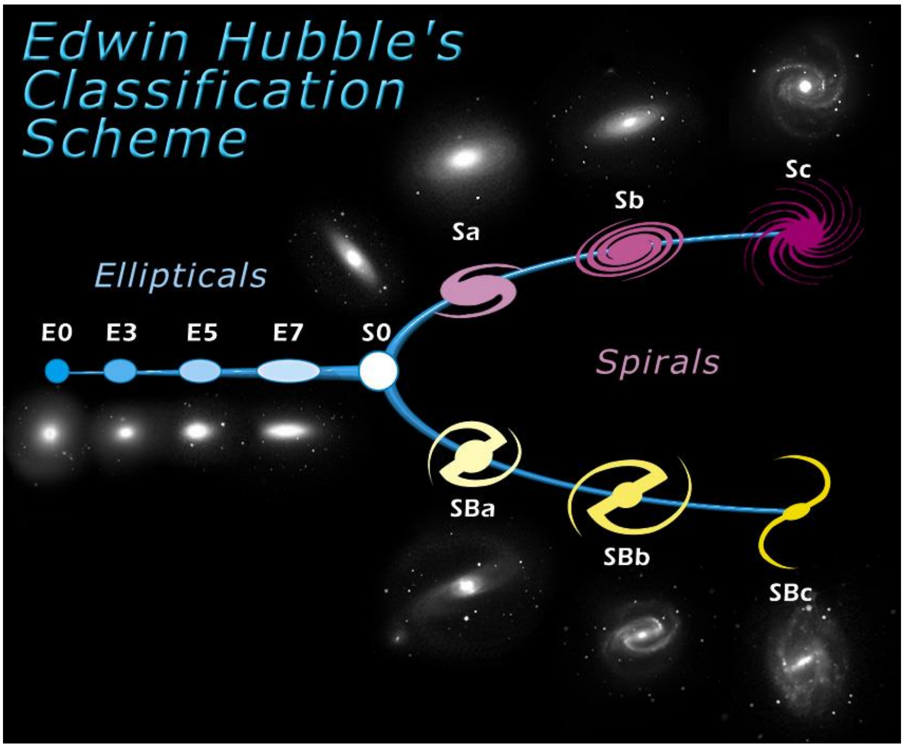


Figure 4.1: Hubble's Tuning Fork. *Hubble classified galaxies according to whether they were elliptical in shape (E0 through to E7, with the galaxy getting flatter in shape moving from left to right), spiral with a bar through the centre (SBa, SBb and SBc), or spiral without a bar (Sa, Sb and Sc). Hubble's classification diagram takes the shape of a tuning fork because he believed that galaxies evolved over time, from left (E0) to right (Sc or SBc): this is now known to be false. (Image taken from Sloan Digital Sky Survey website, http://skyserver.sdss.org/dr1/en/proj/advanced/galaxies/tuningfork.asp (accessed: 28 November 2011)).*

Throughout most of the 20th century, cosmology was hampered because there were insufficient records of individual galaxies to provide robust statistical evidence for its conjectures. Sky surveys, which record the existence of astronomical phenomena, had been conducted since the late 1700s, but even by the mid-to-late twentieth century, the most advanced surveys comprised only thousands of images which were not of a high enough quality to detect the physical features of galaxies necessary for classification, and simply not of a scale to provide answers to cosmological conjectures to a high degree of statistical certainty. Furthermore, because they were expensive for research institutions to procure, many scientists hoping to conduct analysis found them difficult to access, thus slowing down research (Finkbeiner 2010).

In the 1980s, new methods of telescopy were developed which promised the opportunity for sky surveys to expand far beyond thousands of images, to produce images of much higher resolution, and to make them more readily available to researchers. Cosmologists saw this as offering the prospect for providing robust quantitative evidence to support their work (National Research Council 1991), and thus of raising the professional status of their sub-discipline relative to other astronomical sub-disciplines.

One such sky survey was the Sloan Digital Sky Survey (SDSS) (Gunn *et al*. 2006). In 2000, it began operation at the Apache Point Observatory in New Mexico, and recorded spectra for over one million objects. However, despite the public release of the SDSS data, one major barrier remained: how to proceed with the task of classifying such a large number of images.

Prior to Galaxy Zoo, the two alternatives for classifying galaxies were: 1) classifications by professional astronomers; and 2) automated image recognition. However, both of these approaches were found to be lacking in crucial respects as methods for classifying the images produced by the SDSS.

In the case of classifications by professional astronomers, this is a very time-consuming approach. Thus, catalogues of classifications by this method usually number thousands or tens of thousands of galaxies (for instance, see de Vaucouleurs *et al.* 1991; Schawinski *et al*. 2007). A sample of galaxies of this size is not large enough to provide sufficiently robust statistical confirmation of theories about galaxy morphology.

By contrast, the idea behind automated image recognition is that algorithms could classify a very large volume of galaxies in a short timescale. However, developers of algorithms have struggled to devise methods that can accurately classify galaxies in the terms set out in Hubble's tuning fork. The methods devised as of 2007 were deficient in a number of ways: some were heavily dependent on the quality, colour and size of the image used (requiring better quality images that have been produced by sky surveys to date), some were too computationally-intensive , whilst others relied on proxies being used for galaxy classes (e.g. an algorithm classifying spirals and ellipticals according to whether they are blue or red, which results in many spirals being classified as ellipticals and vice versa). Finally, other algorithms have been devised that are more reliable at classification, but they sort according to criteria such as *'asymmetry'*, *'clumpiness'* and *'concentration'* (Conselice 2006, p. 1390), and thus have not been able to provide results that are comprehensible in terms of Hubble's tuning fork, in turn meaning the results of these algorithms are not able to provide statistical evidence for or against proposed cosmological theories (Fortson *et al*. 2011).

4.3 The Galaxy Zoo project

In 2007 there were no methods available that would simultaneously produce classifications of galaxies that were comprehensible to the field of cosmology and be of sufficient scale that would ensure statistical robustness. Galaxy Zoo was developed with the purpose of helping to provide answers to at least some of the major questions in cosmology. This Section gives a brief overview of the initial development of Galaxy Zoo by the project's core team of cosmologists and software engineers. It then describes the work of Galaxy Zoo since it was launched, and its successes in terms of resulting in scientific publications. It will be seen how the project's core team have relied upon the recruitment of both a large number of volunteers to generate the data and groups of astronomers from other institutions to work on this data, producing results that have been disseminated across the academic cosmological community. Then, I will discuss contemporary developments where key members of the core team are attempting to develop the Citizen Science Alliance (CSA).

Through this account, it will be seen that the core team were able to pursue their interests as cosmologists within the contexts presented in Section 4.2. Furthermore, it will be seen that as the previously-unanticipated success of Galaxy Zoo became apparent, new possibilities suggested themselves to the core team, namely that of leading a large and growing movement of Citizen Cyberscience. This, in turn, renewed their interests in securing the future of Galaxy Zoo, despite challenges that they perceive from developments in providing new sources of cosmological data.

4.3.1 The launch of Galaxy Zoo

Galaxy Zoo was launched in 2007 after a doctoral student (Kevin Schawinski) and a postdoctoral researcher (Dr Chris Lintott) developed the concept of recruiting members of the public to classify the galaxy images produced by the SDSS. Schawinski was motivated by his experiences of classifying by hand approximately 50 000 SDSS images (Schawinski *et al*. 2007), and he concluded that it would be impossible for a single astronomer, or even a team of astronomers, to scale up to the full SDSS data set. He and Lintott had been inspired by the examples of other citizen science projects, and wondered whether a similar approach might allow the full SDSS data set to be analysed. Initially, they had intended for volunteers to classify galaxies only according to whether they were spiral or elliptical, but they made contact with another researcher at Oxford, Dr Kate Land, who was interested in the direction in which spiral galaxies rotated (clockwise or anti-clockwise) in order to test the correlation between this rotation and their position in the sky. A decision was made that Galaxy Zoo volunteers should be asked whether an image showed a spiral or elliptical galaxy and, if a spiral, what is its direction of rotation.

During the months preceding the launch of Galaxy Zoo, a number of other individuals were recruited by Lintott and Schawinski to help develop the project. At first, the project did not have any funding; instead, people were recruited voluntarily, and infrastructure (such as the servers) was given as a favour[12]. The project was hosted on servers at John Hopkins University. Software engineers were recruited to design the project website and the volunteer interface, aided by somebody involved in professional astronomy outreach, experts on the Sloan Digital Sky Survey and in galaxy classification were on hand to give advice about the science underpinning the project, and there was also an expert in electronic data archives (Fortson *et al*. 2011).Together, these ten people formed the initial core Galaxy Zoo team[13].

The Galaxy Zoo project was launched in July 2007. This launch included an online article on BBC News, as well as interviews on local and national radio. These were disseminated via global media outlets. Despite this coverage, they were only anticipating a relatively small number of volunteers (in the thousands) and drawn largely from contacts they had within astronomy and enthusiasts with a prior interest in the subject. In the event, however, their expectations were vastly exceeded: by the following April, over 100 000 volunteers had actively contributed to the scientific work of the project. The surprise of the core team about how many volunteers signed up is reflected in the following quotation from one of my interviews:

> **Extract 4.1:** *'The biggest doubt was would there be enough people interested in doing this and what we planned for was a very small number of people, contacts and amateur astronomers, and people we know to spread the word, and we thought it would take about three to five years to get every galaxy viewed a couple of times, and of course we were off by orders of magnitude, the response was so overwhelming when we launched the website, a server physically melted'* (Interview GZ-a/2)

4.3.2 The scientific work of Galaxy Zoo

The scientific work of Galaxy Zoo is directed towards the goal of producing results which can be disseminated to the broader academic community, for instance through publication in peer-reviewed journals and conference presentations. This work has two major components. The first is the core work of the project, which is the work that the project is designed to complete (i.e. the classification of galaxy images and the generation of statistics about these classifications) and is the primary task of volunteers. The other component comprises spin-off projects, which have resulted from volunteers spotting unusual objects in galaxy images and posting them to the forum. Initially, the core team were expecting only core work to be done, and did not anticipate that volunteers would make serendipitous discoveries of new astronomical objects. However, the potential of such discoveries emerging

[12] This comes from my conversations with those involved in Galaxy Zoo.
[13] This core team has been largely stable over time, although a couple have since moved on from the core team, and a software engineer has been employed since early 2009 and is now considered part of the core team.

from the project were seen by members of the core team as opening up new possibilities, including recruiting more volunteers to perform the core work through the often-extensive media coverage of some of these discoveries, and offering a justification for continuation of the project even if and when its core work becomes superseded by automated methods.

The core team of Galaxy Zoo, although responsible for the development and day-to-day running of the project and the production of the Galaxy Zoo data set, do not themselves perform work on the statistical and classifications data produced by the project. Instead, this work is done by a number of groups of research scientists at different institutions who are using the data produced by the work of the project to investigate particular properties of galaxies and the universe. Galaxy Zoo enrolled these groups by offering them the prospect of accessing data which would enable them to answer interesting cosmological questions, which thereby helped to shape their interests through encouraging them to pursue this line of research (rather than other alternatives), and also by the prospect of pursuing opportunities for public engagement. The project has not provided any funding for these scientists. Instead, these groups of scientists had to secure funding for themselves. For example, Dr Karen Masters, a postdoctoral researcher at Portsmouth University, became interested in using the Galaxy Zoo data to study the relationship between galaxy morphology and colour, and secured an early-career grant from the Leverhulme Foundation to fund this work. This grant is not enough for her to be able to work full-time on Galaxy Zoo data and thus she is also involved in other research projects.

The project website Before considering the scientific work of Galaxy Zoo in more detail, I first wish to highlight some key features of the project website (www.galaxyzoo.org (accessed 9 December 2011)). See Figure 4.2): because this is the main interface for interaction between Galaxy Zoo scientists and volunteers, it is the site where much of the project's scientific work is accomplished. This website has a number of sections. These include a page describing the history of the Galaxy Zoo project, explaining what has been learnt so far and containing links to a list of publications and to a page where the publicly-released data set can be downloaded. The website also includes a section through which volunteers can participate in classifying galaxies (including: registration, a tutorial, and the classification interface itself. Subsection 4.3.2 will give more details). There are also pages on the site that aim to educate members of the public about astronomy in general, and about the science that underpins the core work of the project.

In addition, there are two features of the website that have more dynamic content and allow for two-way interaction between scientists and volunteers, and amongst volunteers themselves. The first is the online forum (see Figure 4.3). Although not present at the very beginning of the project, the forum was added approximately two weeks after the project started because the project scientists were unable to handle the flow of emails they were receiving from volunteers on a range of subjects, such as technical support and questions about science, and the scientists hoped that many of these questions could be answered by other volunteers. The forum has proved very popular: it is lively, with dozens of active posters, many of whom contribute on a daily basis. Furthermore, there have been a number of offline meet-ups arranged, with forum members travelling (including from overseas) to attend.

The forum comprises a number of sub-forums, covering a range of topics. Some forums encourage the volunteers to post images of galaxies that have particularly interested them (for instance, if they have unusual features), and the volunteers are able to discuss the images amongst themselves and bring to the attention of project scientists those objects that might be worthy of further investigation: as I will explain below, some of these have led to spin-off projects that have had a great deal of scientific interest. Other sub-forums are themed around discussing science. Others still relate to sharing and solving technical issues that volunteers may be experiencing in participating in the project, whilst the aim of others is to provide a space for volunteers to discuss issues that relate in no way to science.

The other major arena for interaction is the project's blog (see Figure 4.4), which consists of posts made by project scientists (along with the occasional post by a volunteer). Again, the blog was not present at the start of the project, but was added in December 2007. Posts are on a number of themes. The blog is where significant news about the project is shared with volunteers such as: the development of the Galaxy Zoo 2 and Galaxy Zoo Hubble projects; coverage in the popular media; and when papers are published in journals. Other posts have endeavoured to open up particular aspects of the scientific process to volunteers. For instance, there was a series explaining the processes of writing and submitting a scientific article for publication, including the roles of particular individuals and the way in which the peer review process typically unfolded. Another series of articles presented the results of interviews with women involved in astronomy, detailing how they became interested in the subject, and the challenges and opportunities they face in their everyday work. Another major theme of posts is that scientists working on research from Galaxy Zoo data will post regularly on their progress, so that volunteers can follow how their classifications are being processed. The Project Blog is the main theme of Section 5.6 in Chapter 5, and will be discussed in more detail there.

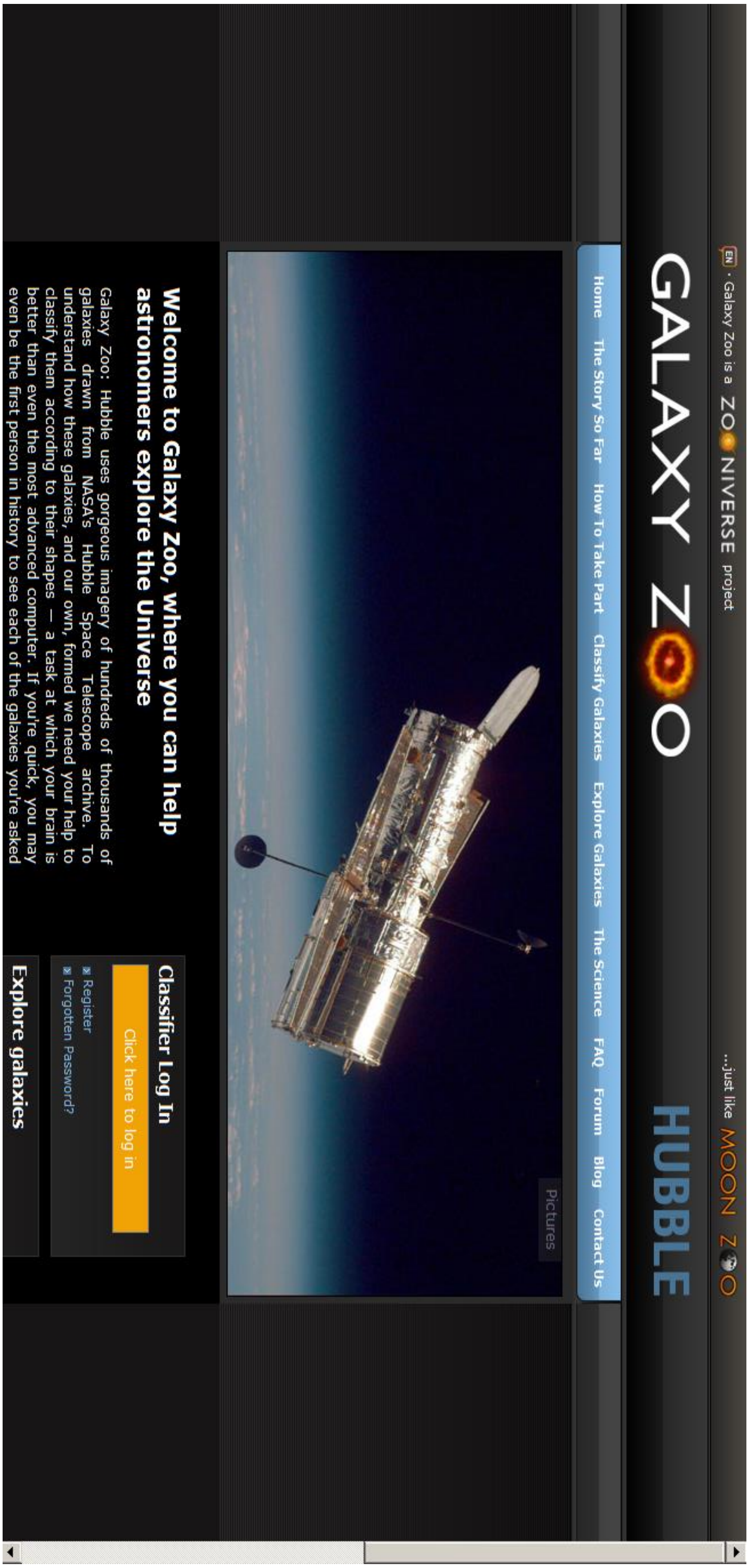


Figure 4.2: The Galaxy Zoo homepage *(http://www.galaxyzoo.org (accessed: 28 November 2011)).*

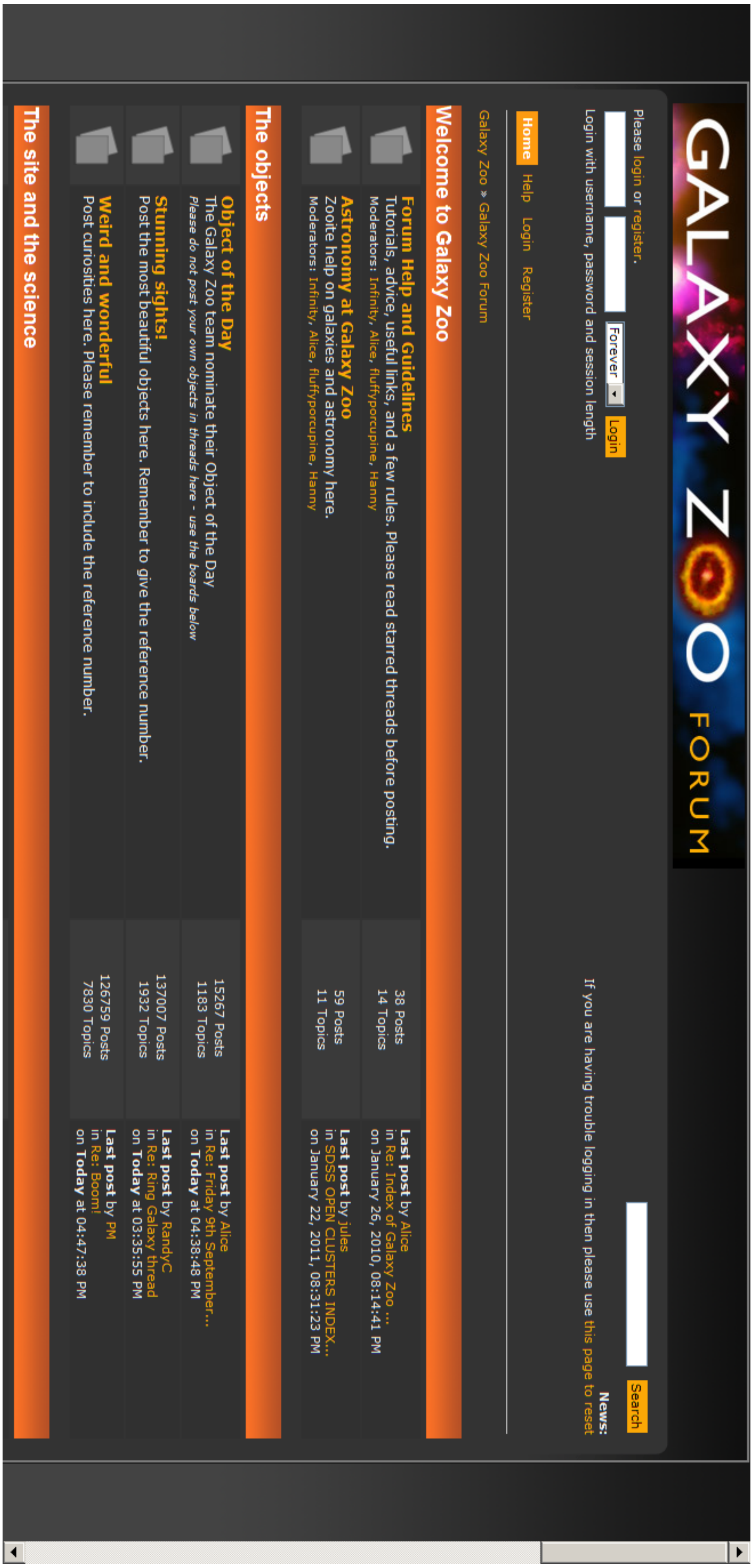


***Figure 4.3: The Galaxy Zoo forum pages**, listing some of the subforums (http://www.galaxyzooforum.org/index.php (accessed: 9 September 2011)).*

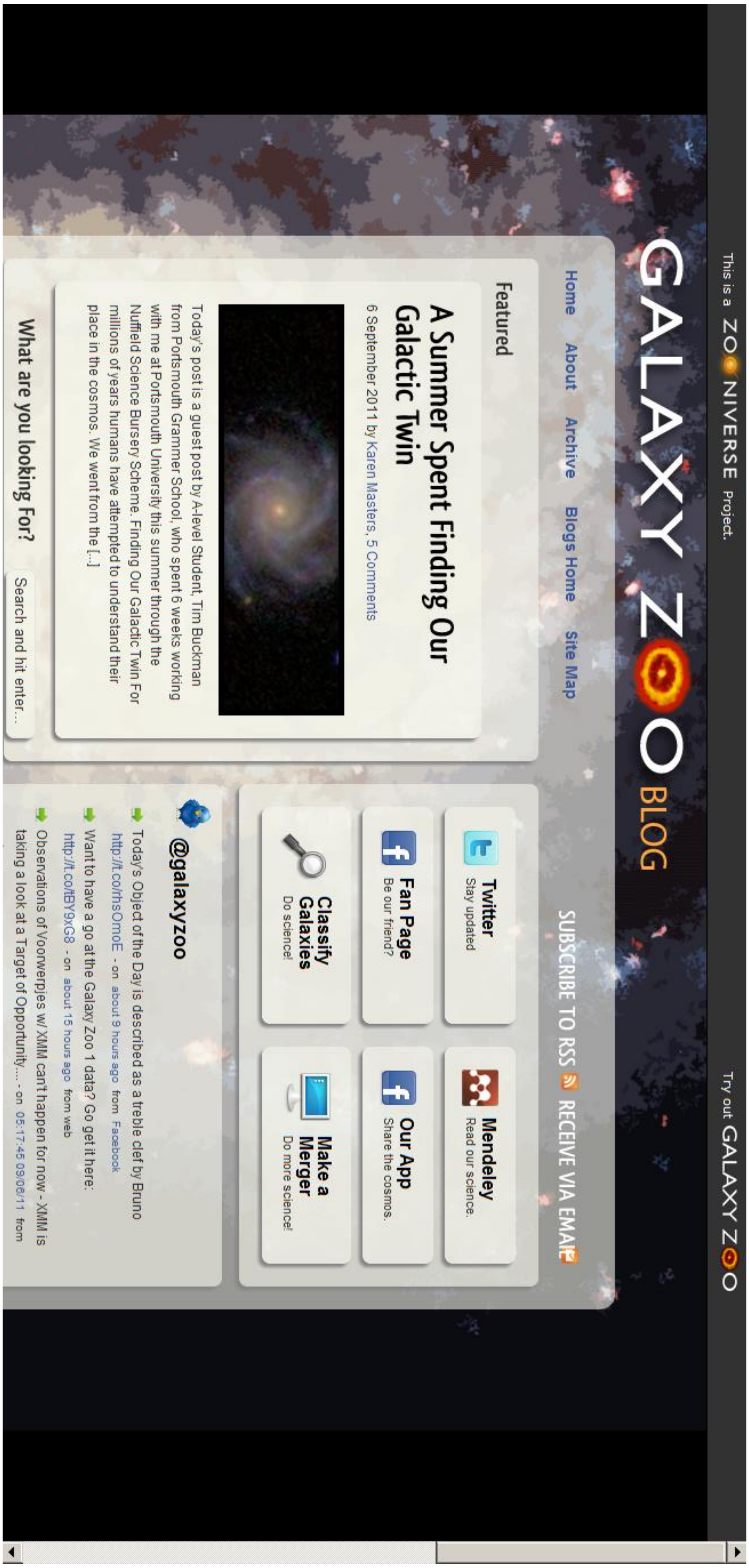


__Figure 4.4: The Galaxy Zoo blog__ (http://blogs.zooniverse.org/galaxyzoo/ (accessed: 9 September 2011)).

The core work: Galaxy Zoos 1 and 2 The first incarnation of Galaxy Zoo ran from July 2007 until February 2009. Its aim was to provide classifications of more than one million images of galaxy candidates (i.e. objects which had been identified as potential galaxies) taken from the SDSS, according to two criteria. The first was whether the image showed an elliptical galaxy, a spiral galaxy, two galaxies in the process of merging, or a non-galaxy (for instance, a star or satellite trail, or if the image was unclear). The second criterion was whether – if the galaxy was a spiral – its arms rotated in a clockwise or anticlockwise direction.

After registering to participate on the project, volunteers were given a brief tutorial in which they were shown example images corresponding to the various categories of galaxies (spiral, elliptical, star/don't know, mergers, clockwise and anti-clockwise). After this tutorial, they were given some examples to practice their classification skills, and were given a short test on a series of images. If their answers were deemed to be correct, they would be allowed to participate in the classification task of the project itself.

Volunteers were presented with a series of images taken at random from the SDSS collection. Volunteers made their classifications using a simple interface (see Figure 4.5). For each galaxy, they were initially presented with four buttons to register whether they believed the image showed a spiral galaxy, an elliptical galaxy, a galaxy merger, or some other object. If they clicked on the spiral button, they were then presented with two further buttons to indicate whether the image showed a galaxy rotating clockwise or anticlockwise. After completing a classification of a galaxy, the volunteer was then asked whether they wished to classify another galaxy. They were able to classify as many, or as few, as they wished in a single session. Volunteers were able to log out at the end of a session, and could log back in at a later date.

During the course of this project, over 100 000 volunteers actively participated. They provided an average of 38 classifications for each of the SDSS images. For each image, the final classification was arrived at by treating each of its individual classifications as a vote – so, for instance, if 28 volunteers said that an image showed a spiral galaxy and 10 clicked on elliptical, it would be classified as a spiral[14].

The resultant Galaxy Zoo data set (Lintott *et al*. 2011) has been used by various selected teams of astronomers to investigate a number of things, including:

- The relationship between the colour of a galaxy and its morphology, to gain a better understanding of how galaxies evolve (for instance, see Bamford *et al*. (2009) and Masters *et al*. (2010));

[14] It should be noted that the process of moving from the volunteer classifications to producing the final data set was far more complicated than this suggests. This is discussed in more detail in subsection 4.4.2.

Figure 4.5: The volunteer interface for the original Galaxy Zoo project
(http://zoo1.galaxyzoo.org/GalaxyAnalysis.aspx (accessed: 9 September 2011).

- The ratio of clockwise to anti-clockwise spiral galaxies in the universe, and also whether there is a correlation between the spin direction of galaxies near to each other (e.g. Slosar *et al*. 2008);
- The properties of merging galaxies (e.g. Darg *et al*. 2010); and
- The relationship between galaxy morphology and black-hole formation in the centre of these galaxies (e.g. Schawinski *et al*. 2010).

Following the unexpected success of the initial Galaxy Zoo project, the core team decided to develop a second Galaxy Zoo project (Galaxy Zoo 2, or GZ2) with the aim of better exploiting the potential of Citizen Cyberscience. It was launched in February 2009, and again received global media coverage. The Galaxy Zoo team had been surprised by the ease with which volunteers handled the original classification system, and, as a result, attempted a more detailed classification scheme in GZ2, using a subset of around one quarter of the SDSS images (these were the brightest images of the original set, the idea being that these images would best allow the volunteers to discern the more detailed features for classification).

In GZ2, volunteers were asked to assess whether: the image showed a galaxy that was smooth and rounded or had other features, such as a disk; whether – in the case of smooth galaxies – the galaxy was round or a more elliptical shape; whether the galaxy was seen edge-on (galaxies generally have a disk shape); whether the galaxy had a visible bulge at its centre, and if so, what was the bulge's shape and to what extent did it dominate the image of the galaxy; if the galaxy is a spiral, and if so, how tightly the spiral arms are woven around the galaxy's centre and how many spiral arms there are; whether the galaxy has a bar at its centre; and finally, whether the volunteer noticed anything unusual about the galaxy, and if so, what it was.

As with the original project, GZ2 was highly successful in recruiting volunteers, and received approximately 60 million classifications until the project ended in April 2010 (Fortson *et al*. 2011). This has produced a data set which is currently being used by a large number of teams investigating the relationships between galaxy morphology and colour, and some of their results are starting to be published (e.g. Masters *et al*. (2011) investigates the connection of bars, bulges and colour).

After the end of GZ2, the Galaxy Zoo team launched a third incarnation of the project, (*Galaxy Zoo Hubble*). They have also been developing and running a number of other citizen science projects in this period: these, along with Galaxy Zoo Hubble comprise the *Zooniverse* under the auspices of the Citizen Science Alliance, and this is described in more detail in subsection 4.3.3.

Serendipitous discoveries In addition to the core scientific work, the Galaxy Zoo projects have also given rise to unanticipated discoveries and spin-off projects, often as a result of volunteers noticing unusual objects in galaxy images and posting them on the forums. Over time, these have become important for the project. One reason is that they increase the opportunity for the project to become associated with novel science that is of interest to the broader discipline. The second is that they have proven useful in helping the project to pursue its core objectives. This is because the idea of a non-scientist being the discoverer of a novel object has proven to be an inspiring narrative, attracting substantial media interest which has helped the project to recruit more volunteers to perform the task of galaxy classification.

Two of these projects will be briefly highlighted here. The first is the discovery of a new object by a Dutch schoolteacher, Hanny van Arkel, who noticed an unusual object in a galaxy image she had been asked to classify and posted a message on the project's online forums in August 2007, shortly after the project was launched (see Figure 4.6). This object became the subject of further investigation, and time on a number of telescopes, including the Hubble Space Telescope, was secured for observing the object.

This has led not only to a number of papers, on which van Arkel has shared co-authorship (for instance, see Lintott *et al*. (2009)), but also significant coverage and recognition for van Arkel. For instance, she has been interviewed for many media outlets, including popular science magazines, as well as more general outlets such as chat shows. She is regularly invited to give talks about her experiences with Galaxy Zoo to schools and other groups of interested lay people, both inside her native Netherlands and throughout Europe, and has also set up a personal website, where she blogs frequently and promotes the Galaxy Zoo project[15]. Finally, in August 2010, a comic book was released for sale which recounts the story of the discovery of Hanny's Voorwerp and the work of the scientists in trying to determine its nature.

The second spin-off that will be highlighted here is the Green Peas project. Here, a volunteer posted a galaxy image on the project forums because they had noticed there was a small green circle on the image and asked what it might represent. This image caught the imagination of the forum users and was labelled a *green pea* by a volunteer and a number of galaxy images featuring similar objects were posted by other volunteers on the forum thread[16]. These objects were investigated further by members of the forum (for instance, by going to the Sloan website and looking at data attached to each image) before finally being studied further by a doctoral student at Yale University. This resulted in a paper being published (Caradamone *et al*. 2009).

[15] http://www.hannysvoorwerp.com (accessed 9 December 2011).
[16] http://www.galaxyzooforum.org/index.php?topic=3638.0 (accessed 9 December 2011).

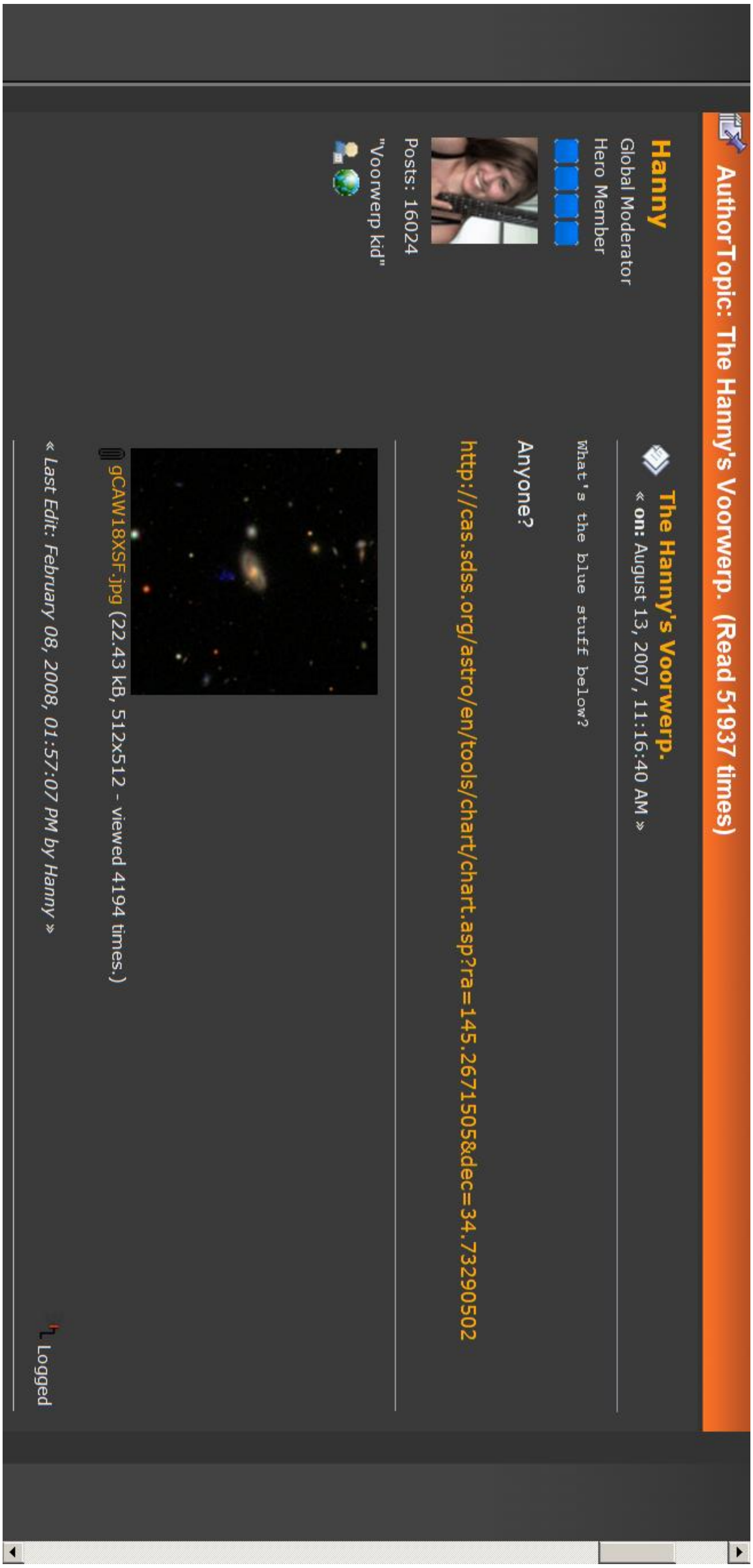


Figure 4.6: The original post by Hanny van Arkel, *in which she posted the image containing the unusual object on the online project forum (http://www.galaxyzooforum.org/index.php?topic=3802.0 (accessed: 9 September 2011)).*

In summary, then, the core team members of Galaxy Zoo have identified, and proven successful in terms of responding to, the broader contexts in which the project is situated, with a number of papers already resulting, and more currently being produced. A critical part of this success has been the core team's ability to interest and enrol a number of groups of people, in particular

- Members of the public to classify the galaxies; and
- Teams of astronomers from other institutions to use this data in their own research, who have been critical in turning the data produced by the volunteers into results that help to advance the state of cosmological knowledge and are of interest to the broader cosmological community. These astronomers, in turn, need to enlist the support of funding bodies as Galaxy Zoo itself is not able to provide them with the financial support to embark on this research.

Buoyed by the success of Galaxy Zoos 1 and 2, the core team set about developing the third incarnation, Galaxy Zoo Hubble. However, because their experiences had made them aware of the potential contributions of Citizen Cyberscience to a range of scientific questions, they no longer sought to restrict their activities to cosmology and instead set about the development of the Zooniverse, a range of projects covering many scientific disciplines, under the auspices of the Citizen Science Alliance. It is to this that I now turn.

4.3.3 The Zooniverse and the Citizen Science Alliance

In this subsection, I will discuss the development of the Zooniverse and the Citizen Science Alliance (CSA). In particular, I will argue that – through the CSA – many of the Galaxy Zoo core team members have developed an interest in becoming the leaders of a large and growing Citizen Cyberscience movement. The CSA is actively seeking to enlist teams of scientists who are interested in setting up CCPs: in exchange for offering technical expertise and a greater ability to promote their project to volunteers, they will be expected to adhere to the norms set out by the CSA. In light of this trade-off and competition from other potential umbrella organizations for CCPs, the CSA's future growth is not certain. In this context, I will explain why the Galaxy Zoo team members see the continued success of Galaxy Zoo as vitally important for the CSA. In other words, not only has the success of the Galaxy Zoo project enabled the core team to pursue their interests as cosmologists within cosmology, but in the future, it has led to an interest in becoming leaders of a Citizen Cyberscience movement: this point will be important in Section 4.5 where I will consider how these individuals see potential threats from contemporary developments in cosmology to the continued scientific success of Galaxy Zoo.

The Citizen Science Alliance (CSA) was set up in 2009, headed by a number of the key members of the Galaxy Zoo core team. Its stated mission is to *'develop, manage and utilise internet-based citizen science projects in order to further science itself, and the public understanding of both science and of the scientific process'* (see http://www.citizensciencealliance.org/index.html (accessed 9 December 2011)). There are currently six partner institutions involved (University of Oxford, Nottingham University, Adler Planetarium, University of Minnesota, National Maritime Museum at Greenwich, and John Hopkins University). The CSA was not planned, or even conceived of, at the start of Galaxy Zoo, but its success led to the realisation by the project's core team that they could use the lessons learned to help other CCPs, perhaps in other disciplines, and exploit Galaxy Zoo's media profile to help these projects with volunteer recruitment.

The CSA oversees the running of the Zooniverse, which was developed with the aims of both developing new citizen science projects within the field of astronomy and also using the Galaxy Zoo platform as a basis for setting up projects in other fields, such as climate science. The basic concept is that scientists from any discipline who have a large amount of data to be processed might approach the CSA and request help in setting up and running a CCP. In turn, if the project is approved, software engineers working with the CSA will help to develop and maintain a project website, including online forums, a project blog and an interface through which volunteers can access and analyse the data. This project will then become part of the Zooniverse, and be overseen by the CSA.

To date, ten projects are part of the Zooniverse, with over 500 000 registered volunteers in total (source: www.zooniverse.org (accessed 9 December 2011)). These projects are:

- ***Galaxy Zoo Hubble*.** This is often referred to informally as *Galaxy Zoo 3*. It asks volunteers to perform tasks similar to those in Galaxy Zoos 1 and 2, namely the classification of images of galaxies, in this case generated by the Hubble Space Telescope;
- ***Galaxy Zoo Mergers*.** This project investigates the properties of galaxies that are in the process of merging by asking volunteers to compare images of such galaxies with computer simulations of merging galaxies. The core team of this project is comprised of scientists and developers from the main Galaxy Zoo project;
- ***Galaxy Zoo Supernovae*.** This project attempts to identify a particular class of star called supernovae in telescope images;
- ***Milky Way Project*.** This project aims to further understandings of how stars develop in the Milky Way galaxy;

- ***Solar Storm Watch.*** This is a collaboration with a team from the Royal Observatory in Greenwich, and asks volunteers to detect and track solar storms from images of the Sun;
- ***MoonZoo.*** This project presents volunteers with a series of detailed images of the Moon's surface from NASA's Lunar Reconnaissance Orbiter and asks them to identify craters, with the goal of developing a better understanding of the Moon's history;
- ***PlanetHunters,*** which uses images from NASA's Kepler mission, which is a search for habitable planets, and asks volunteers to examine these images to identify and mark potential new planets;
- ***Old Weather.*** This is a climate science-driven project. It asks volunteers to transcribe handwritten entries from ships logbooks from various points in history and around the globe with the aim of studying historic changes in global climate;
- ***Ancient Lives.*** This is a collaboration with a team of papyrologists, and asks volunteers to transcribe fragments of ancient Egyptian papyri;
- ***Icehunters.*** This project presents volunteers with telescope images and asks them to identify objects in the Kuiper Belt, which is a ring of chunks of ice lying in the Solar System beyond Pluto. In 2015, the New Horizons space probe will be flying past Pluto, and the aim is to identify Kuiper Belt objects that the probe might photograph.

The CSA is actively canvassing for more researchers to apply to set up and run projects under its auspices. In order to interest these researchers in running such a project, it has run conferences, bringing together researchers (including those interested in setting up a project), funding bodies, and volunteers to discuss and map out the future of citizen science projects and explain to researchers how Citizen Cyberscience might afford possibilities in answering important scientific questions in their own discipline. In order to assess and select which projects it may help, the CSA has a long questionnaire which a team of scientists interested in setting up a project needs to fill in, and a subsequent, rigorous assessment process.

The benefits to a team of scientists who have become interested in running a CCP are clear. Joining the CSA would allow them to set up their project based upon an already-proven framework. It gives them access to the expertise of Galaxy Zoo core team members across a number of domains, including technical (such as designing the volunteer interface), project promotion through the media, and how to recruit and interact with volunteers (for instance, frequency and content of feedback in the project blog). It also offers the benefit of ready access to a large pool of active volunteers (once a member of the public signs up to one Zooniverse project, they are automatically signed up to all projects and these are cross-promoted), as well as the promotional advantages of being associated with an already-strong brand. A final advantage is that involvement with the Zooniverse may endow a project's results added scientific credibility by its association with citizen science projects that have already won credibility (see subsection 4.4.2).

In order to recruit interested scientists, the Galaxy Zoo project is regarded by key individuals as very valuable for the CSA. It is very much the Zooniverse's flagship project (in its current incarnation as Galaxy Zoo Hubble), both in the sense that it is held up by the CSA as the example of what citizen science makes possible (and is thus used to attract the interest of scientists who might seek to set up such a project), and it is the main project through which volunteers are recruited to the Zooniverse (and, through this, can be introduced to other projects). Compared to the other Zooniverse projects, Galaxy Zoo has:

- **A larger number of volunteers than other Zooniverse projects**, as is explained in the following quotation from one of my interviews with a core team member:

  > **Extract 4.2:** *'In terms of the number of people that come back to the project every day, it's still our busiest project, it does have the biggest community'* (Interview GZ-a/1);

- **A more active and vibrant volunteer community.** The Galaxy Zoo forums have a much larger number of registered users, who post at a higher rate on a wider range of topics (including topics completely unrelated to science), than the other projects. Furthermore, the Galaxy Zoo community organises regular off-line meetups, which is largely absent from the other projects;
- **Higher impact on, and acceptance by, the broader scientific community,** measured both in terms of the number of scientific articles published and the number of citations accrued by these articles (to date, Galaxy Zoo has published 24 articles in peer-reviewed journal articles, whilst the other projects have published only preliminary results so far, if at all);
- **Far greater media coverage.** Galaxy Zoo enjoys much broader coverage in mainstream media outlets than the other projects. Media coverage has proven invaluable in terms of recruitment (as mentioned in subsection 4.3.1) for Galaxy Zoo.

In summary, then, the CSA can therefore be understood as seeking to become a gatekeeper for, and to place itself at the head of a movement of, online citizen science. In order to enrol potential collaborators, it uses the Galaxy Zoo project as a flagship to interest scientists in the potential of Citizen Cyberscience and to join the Zooniverse. Thus, it can be understood that the Galaxy Zoo core team view the continued success of the project as important in pursuing their more recently-developed interests within the context of Citizen Cyberscience. In light of the discussion in subsection 4.3.2, they are keen to ensure the continued engagement of both volunteers and astronomers interested in working with the Galaxy Zoo data.

4.4 Enrolling other astronomers

In their study of Galaxy Zoo volunteers (which has been co-authored with members of the Galaxy Zoo core team), Raddick *et al*. (2010) concluded that the overriding factor motivating volunteers was the idea that they contribute to scientific progress through their continued participation in the project. This has been strongly echoed by my study of climateprediction.net, as well as in my own observations of, and conversations with, Galaxy Zoo volunteers. Thus, I will not problematize for the moment the question of how the project's core team viewed the challenge of how to retain volunteers, and I will assume that they saw this in terms of how to reassure volunteers that they were indeed assisting the advancement of scientific knowledge.

Instead, in this Section, I consider the problem of the enrolment of the astronomers who might work with the Galaxy Zoo data, and in particular how it is perceived by the core team members. The purpose is to highlight some of the issues that the Galaxy Zoo core team both have had to address already to succeed in this, and also those they believe they will need to continually address if they are to continue this success.

I will first present the case that, although initially attracted to the project for both scientific reasons and the opportunity of public engagement in science, only the former (in the sense of providing data of a scale and quality unmatched by other sources that will help them to produce research that enables them to secure recognition in the field of cosmology) provides a major motivation for astronomers to continue their collaboration with the core Galaxy Zoo team. Through their experiences with the project, these astronomers seem to have concluded that the project affords them few opportunities to engage with the public in the manner they desire and, because they often have other avenues for pursuing this readily available to them, the public engagement element is therefore not sufficient (nor, indeed, necessary) to retain the enrolment of these astronomers: instead, it is all about the data. I will then turn my attention to the issues that the Galaxy Zoo core team have had to address in order to secure the credibility of the data generated by the project.

4.4.1 Motivation for astronomers to become enrolled by Galaxy Zoo

Here, I will argue that the interest of astronomers from outside the core team in using the Galaxy Zoo data has depended upon the project's ability to produce data of a scale and quality unmatched by other sources. My basic argument is that whilst these astronomers have come to view Galaxy Zoo as an *obligatory passage point* from the point of view of their scientific work[17], from the point of view of their own particular public engagement objectives

[17] The term *'obligatory passage point'* was coined by Callon (1986a, p. 199) and refers to a social actor with whom another social actor **must** collaborate for this latter actor to pursue a particular goal. In the case here, Galaxy Zoo is an obligatory passage point for

many have concluded from their experiences in the project that the project is not an obligatory passage point for pursuing these goals as there are other public engagement activities readily available to them that they believe equally, or better, allow them to pursue these[18]. For the sake of clarity, I wish to emphasize that this subsection focuses only on those astronomers from outside the Galaxy Zoo core team: indeed, those from within the Galaxy Zoo core team appear to have a very different perceptions of the project's potential for public engagement and often spoke enthusiastically about opportunities for improving volunteers' understandings of the scientific process in my interviews with them[19].

The following quotation is typical of the interviewees' attitudes regarding the relative importance of Galaxy Zoo enabling them to conduct science and public engagement:

> **Extract 4.4:** *'Bob Nicol[20] suggested I do an analysis similar to what I had done before but with these Galaxy Zoo morphologies…my motivation for getting involved was mainly scientific…it was nice to actually do something that involved the public so I did appreciate that aspect to it, but at the time I was just excited with this huge amazing dataset' (Interview GZ-b/4)*

Here, it can be seen that the primary motivation for becoming involved was that it fitted in well with the interviewee's existing research interests, and the large dataset offered the prospect of advancing these interests: without this, it is clear that the public engagement element, though appealing to this scientist, would not be enough to secure their interest.

In order to better understand why, it is necessary to consider the particular objectives that motivate them to engage with the public, and the extent to which they believe avenues other than Galaxy Zoo enable them to pursue these objectives. My interviews suggested a scientist might have at least one of four underlying motivations for engaging with the public. One motivation is that some interviewees expressed a desire to encourage more children, especially girls and ethnic minorities, to consider careers in science. They see public engagement as a means to break down stereotypes about the sorts of people who embark on scientific careers, and to suggest to schoolchildren that they should not be deterred from considering such a career on the basis of gender or ethnicity. This is exemplified by the following interview extract:

> **Extract 4.5:** *'I am very interested in getting all different sorts of people involved: that includes women, minorities. I think science has such a hierarchical structure and uniform image of what it's like to be a*

---

the astronomers in the sense that if these astronomers wish to pursue large-scale statistical analyses of galaxies then they must engage with the Galaxy Zoo project as this is currently the only source of data available which would allow them to do this.

[18] During the interviews, the astronomers cited a wide range of public engagement activities in which they already participate, such as talks (in schools, to amateur astronomer groups, public lectures etc.), open days at their laboratory or observatory (including those aimed at children only), courses for schoolchildren where they are given the opportunity to undertake projects in the astronomers' laboratories, personal blogs or Twitter accounts reporting their research, and writing for popular science magazines.

[19] For instance:

> **Extract 4.3:** *'This is a great, great way to show people how scientists actually work, how science works, how we form hypotheses, how we get data, how we analyse it, how we write up a paper and the whole refereeing process, and we had a great thing saying,* "Follow the story of your discovery from beginning through to interpretation." *'* (Interview GZ-a/2)

[20] Bob Nicol is a core team member, based at the University of Portsmouth.

scientist...I think that turns away a lot of people when they could be very interested in it, so a lot of what I do is try to break that stereotype' (Interview GZ-b/2)

In this case, however, they found Galaxy Zoo to be lacking in this respect compared with other alternatives. This is because, drawing on their own experiences of when they themselves were inspired to pursue a scientific career, the interviewees stated that they believed that focussing efforts on young children would be the most effective way of achieving their objectives, and that Galaxy Zoo is inappropriate for this age group:

Extract 4.6: *'I tend to focus on children…I don't think many children would do Galaxy Zoo…I feel like I have the biggest impact if I talk to children because I think the biggest impact is if they see me as a scientist, and presumably I break any stereotypes they have about scientists…There have been studies about stereotypes that kids have about scientists'* (Interview GZ-b/1)

A second reason is that involvement with public engagement activities can improve a scientist's career prospects. As discussed in the Introduction to this thesis, there is an increasing culture that expects scientists to participate in public engagement activities, and taking part in Galaxy Zoo is a way of demonstrating a willingness to do so. This is illustrated by the following interview extract:

Extract 4.7: *'I've been applying for faculty positions in the US and in the US, there's science of course, there's research, there's teaching, but then there's also a public outreach component so I think they like that, that I've been involved in public outreach…Galaxy Zoo was one of the things that I think personally has been helpful in my career because I've done a lot of public outreach so it's good to show that I'm continuing with that'* (Interview GZ-b/3)

Again, however, this goal could be pursued via other methods public engagement, for instance through involvement in one of a number of formal initiatives funded either publicly or by charitable foundations, which would also demonstrate an astronomer's commitment to public outreach to a potential employer.

The third motivation is that some scientists reported that they feel a responsibility to keep members of the public informed about their work, because a large part of this work is publicly-funded and thus the public should be reassured their money isn't being wasted, as explained by the following interview extract:

Extract 4.8: *'I'm being paid by taxpayers' money and I feel like it's almost my duty to pay some of that back to the community'* (Interview GZ-b/3)

A fourth reason is enjoyment about sharing something they are passionate about:

Extract 4.9: *'I really love the enthusiasm of the public and the response to astronomy really, so I find it very rewarding to interact with the public'* (Interview GZ-b/3)

For these third and fourth motivations, Galaxy Zoo may seem like an attractive way of pursuing public engagement. However, a number of the scientists with these motivations expressed frustration or scepticism about the opportunities for public engagement afforded by the project and suggested that they were better able to pursue their public outreach goals via other methods. Some felt that opportunities to communicate through Galaxy Zoo with the public were limited to a small number of the project scientists, for instance:

> **Extract 4.10:** *'I felt like there's not much new I can add to what…some of the main characters are already writing on the blog'* (Interview GZ-b/3))

Others had found that, in spite of the number of members of the public who had enrolled, there was only meaningful engagement with a very small subset of volunteers, for instance:

> **Extract 4.11:** *'We don't get a lot of response…when I first wrote blogposts, I though there were 400 000 people in Galaxy Zoo and so I was writing for 400 000 people, but at the end of the day there's far fewer people reading it, which is quite disappointing'* (Interview GZ-b/1)

Likewise, there was little expectation that many volunteers, beyond a *'hard core'*, follow scientific threads on the forum, instead seeing the forum as *'a place to hang out with friends and talk about general things rather than discuss science'* and concluding that the vast majority of volunteers are only *'tangentially interested in astronomy'* (Interview GZ-b/5) (**Extract 4.12**). A number of these scientists have concluded that engagement through Galaxy Zoo is superficial compared with other methods of public engagement:

> **Extract 4.13:** *'The public outreach I was doing before was…going out and talking to people about science and doing workshops with them, we were doing workshops with kids and stargazing at night, and that was really rewarding because you were doing it on a personal level…the Galaxy Zoo outreach is slightly different because…the context is slightly shallower'* (Interview GZ-b/6)

Thus, although Galaxy Zoo apparently offered the prospect of reaching a large number of people, and indeed many of the scientists had reported enjoying sharing their work with lay audiences, those interviewed often expressed the view that they had only found a deep engagement with a small number of volunteers and thus the project did not afford a possibility to truly engage with a broad mass of laypeople.

To summarise: although many of the astronomers from outside the core team were attracted to Galaxy Zoo because they had a prior interest in taking part in public engagement activities, the project needs to have a very clear scientific justification in order to continue to interest them because they found that the project didn't really afford new opportunities to pursue their particular public engagement objectives.

4.4.2 Gaining credibility in the broader scientific community

The Galaxy Zoo core team members have worked hard to secure the credibility of their project's output. This has been considered important in order to secure enrolment of astronomers to use the project data because:

1) These astronomers, in turn, need to enlist the support of funding bodies: to achieve this, these bodies need to be convinced of the credibility of the Galaxy Zoo data;
2) These astronomers are seeking to gain recognition of their peers, for instance through conference presentations or through the publication of journal articles that are read and cited by other astronomers. Such recognition depends upon the wider credibility of Galaxy Zoo data.

This subsection discusses how this credibility was achieved: it has been hard-won in the face of some scepticism, and the many painstaking steps taken to do this bears testament to how important and difficult this has been for the project's core team. As will be seen, these steps were not just a case of providing technical solutions (such as writing software or introducing extra steps into the processing of data) but also involved management of volunteers, and thus involved social learning on the part of the core team about the community dynamics that formed around the project.

The data produced by Galaxy Zoo is in competition with other alternative methods of data generation, namely automated image recognition and classifications produced by professional astronomers, and core team members were made aware that those involved in such methods were quite resistant to a Citizen Cyberscience method claiming to be superior and thus threatening to render their work irrelevant. For instance, it was reported by one of my interviewees that one such individual had told him *'Galaxy Zoo will put [us] out of jobs'* (Interview GZ-a/6) (**Extract 4.14**).

Professional astronomers might also be expected to be keen to preserve the boundary between science and the rest of society (see Chapter 2, subsection 2.2.1), and thus be unsympathetic to a project such as Galaxy Zoo which takes the generation of data out of the hands of professional scientists and into the realm of amateurs. In this respect, a number of my interviewees reported that many astronomers outside of Galaxy Zoo expressed scepticism about the quality of the data produced by the project. The following is typical of the experiences reported by interviewees when they presented Galaxy Zoo to other astronomers:

> **Extract 4.15:** *'The first few talks I gave, I would get a lot of questions saying you can't trust these classifications'* (Interview GZ-b/1)

In particular, core team members became aware of three issues that were raised to cast doubt on the integrity of Galaxy Zoo data. The first was whether members of the public had sufficient knowledge and experience that their classifications would be reliable, as is explained in this quotation:

> **Extract 4.16:** *'There were people who have spent years of their lives learning how to classify galaxies and how these classifications fit in the context of galaxy evolution, and so they're sceptical that some supposedly random person could just click a few buttons'* (Interview GZ-b/4)

One way in which it was suggested that this would cause a problem was that volunteers' classifications would change considerably as they became more experienced at the task; on the other hand, experienced professional astronomers could be relied upon to provide results consistent over time as their skills were already honed. For example, in the Sloan Digital Sky Survey images, most spiral galaxies are red whilst most elliptical galaxies are blue but there is some crossover (blue spirals and red ellipticals): some concern was expressed that the public, lacking sufficient knowledge, would not realise that such crossover existed and would get into the habit of simply classifying by colour (i.e. all red images as spirals, all blue images as ellipticals). This might mean that Galaxy Zoo data would compare less favourably both to data produced by expert classifiers in possession of the necessary knowledge, and also compared to automated methods, which might be designed to ignore colour and classify simply according to shape.

Another issue raised involved doubts that members of the public could be relied upon to report their classifications honestly: many astronomers wondered aloud that whilst professional astronomers could be trusted because, it was argued, they had the interests of the discipline at heart, would members of the public have less savoury motives? Again, such a concern was seen by the core team as threatening the credibility of the Galaxy Zoo data relative to data produced by other methods.

A final issue regarding the Galaxy Zoo method in relation to automated methods is that humans may have cognitive biases when classifying, which would systematically skew the results: for instance, if, for a particular feature, it is unclear from an image which category it fits into in the classification theme, it may be that humans are more likely to choose one of the categories over the alternatives (e.g. if it is unclear whether a spiral galaxy rotates clockwise or anticlockwise, is there a bias in favour of one over the other?).

In order to address these concerns, the Galaxy Zoo team has taken a number of steps to secure its reputation as a source of credible data (Fortson *et al*. 2011). Some of these steps have been straightforward, but some have involved more painstaking work over a longer period of time, thus demonstrating both the extent to which the Galaxy Zoo team perceived the threat to the credibility of the project's data, and the importance they attached to overcoming this. The following quotation from an astronomer who was part of the core team suggests that the efforts to secure the data's credibility were costly for the project:

> **Extract 4.17:** *'I think it's questionable whether [the Galaxy Zoo method of classifying galaxies] works out in the sense that, yeah, you get a lot of things for free but…you also have to calibrate them out, which turns out not to be so simple any more'* (Interview GZ-a/5)

The more straightforward measures were aimed at addressing scepticism that volunteers could be trusted not to make malicious classifications, and whether a lack of expertise and experience on the part of volunteers might compromise the quality of their classifications. One such measure was identifying and removing data from volunteers that was clearly out of line with other volunteers' data (and thus suspected as being malicious), the purpose of this is to reduce the influence of malicious classifications on the project's data.

Another relatively straightforward measure was introducing a weighting system for volunteers, with the classifications of volunteers who most closely match those of the other volunteers being assigned a higher weighting (Fortson *et al*. 2011), the intention being to reduce the influence of malicious users whose classifications had not been identified and removed, and also to increase the influence on the data of those volunteers who are more skilled at classifications and thus counter criticisms that the data would be compromised by volunteers with less-refined skills.

A third measure was to compare the results of volunteer classifications with results generated by other methods, in order to increase their credibility by showing that they agree with already-credible methods (this is presented in an early Galaxy Zoo paper (Lintott *et al*. 2008), with results suggesting a rate of agreement of well in excess of 90% between Galaxy Zoo classifications and those by professional astronomers). The following extract from an interview with a core Galaxy Zoo team member suggests this had some success in securing the credibility of the project's data:

> **Extract 4.18:** *'When Galaxy Zoo results would fit into, be consistent with other results but with more statistical power, I would get fewer sceptical comments'* (Interview GZ-b/4)

A more onerous task was to mitigate scepticism that human bias would compromise the Galaxy Zoo data. This task had a number of dimensions: not only did it involve technical aspects (in terms of devising and implemented methods of testing bias) but also social factors (in terms of volunteer management).

To date, there have been three bias tests. The first was to test whether the colour of an image of a galaxy (red/blue) had any impact on whether it was classified as a spiral or elliptical: to test this, black & white images of galaxies were included alongside colour images in the pictures presented to volunteers for. The task was further complicated when it was found that volunteers changed their classification habits during the period of the bias testing, in the sense that they tended to take more time over each classification, posing an additional challenge to the Galaxy Zoo scientists. This was dealt with, and the final conclusion was that there was little difference between the classification of full-colour and black & white images, which helped in countering claims that a lack of expertise and experience on the part of volunteers might mean that they would classify according to colour rather than shape (Forston *et al*. 2011). The second bias test was to see whether there was a bias in favour of classifying spirals as clockwise rather than anticlockwise (or vice versa) in the cases where galaxy images were blurry or faint. In order to test this, a series of mirror images of galaxies were included in the galaxies presented for classification and a bias was indeed detected. This allowed for the bias to be corrected, thus enhancing the credibility of the Galaxy Zoo data (Land *et al*. 2008).

The final bias test was driven by scientific work to ascertain the properties of a subgroup of galaxies: those with *Active Galactic Nuclei* (AGN), which are compact regions at the middle of a galaxy. AGNs shine very brightly in the middle of galaxy images, and can thus obscure most of the galaxy, and this bias test was conducted to assess the impact of the presence (or otherwise) of an AGN on the classification of the galaxy's morphology.

This test involved simulating images of AGNs, superimposing them on galaxy images and then including these images in the set presented to volunteers for classifications. This allowed for comparisons of the classifications of a galaxy image with and without a simulated AGN superimposed on it.

The example of this third bias test is interesting. It involved a great deal of technical and statistical work in terms of conceiving the test, coding software to simulate images of AGNs and combining with galaxy images, and inserting the images in the volunteer interface. However, it also involved a great deal of additional work in terms of managing the volunteers themselves. The following occurred: a post was made on the Galaxy Zoo blog briefly explaining the test, which was met with a number of strongly-worded responses in the comments section from Galaxy Zoo volunteers. Some expressed anger that they believed it meant they were not doing real science (for instance, one volunteer wrote: *'We're trying to help out here with some boring science work, we're not stupid guinea-pigs, clicking away on fake pictures'*, even to the extent of accusing the scientists involved of *'scientific malpractice'*
(http://blogs.zooniverse.org/galaxyzoo/2011/01/11/fake-agn-galaxies/ (accessed 9 December 2011)).

As a result, a number of volunteers threatened to leave the project. In response, the scientists involved had to respond to volunteers' criticisms in the comments section of the original blogpost, devise new blogposts to assuage the volunteers, and also communicate with certain individuals who were seen to be leading the opinions of others. One scientist recounts their experience:

> **Extract 4.19:** *'The person who kicked up a fuss was a complete pain, with no interest in trying to understand our explanations…It takes one person to say,* "This is ridiculous, what you're doing is immoral," *for 100 to line up behind them…I have sent [this person] numerous personal emails trying to explain to him why we're doing this particular thing with the project'* (Interview GZ-a/1)

In other words, in addition to the technical challenges of running this bias test, the scientists also had additional workloads managing volunteers. As the above quotation suggests, this involved more than simply writing blogposts, it also involved the scientists having to make strategic moves and develop an understanding of the dynamics amongst volunteers in terms of being able to identify and target those volunteers who are particularly influential in terms of other volunteers' opinions.

A final, important, way of demonstrating the credibility of the data in the face of scepticism about human bias was that the project could point to the very large number of volunteers who had classified the galaxies. The robustness of this data rests, to a large extent, on the fact that the classifications have been performed by such a large number of people. The ideal would have been for each classification to be performed by a different person, in order to ensure that every classification was independent of the others. In other words, it was more desirable for a very large number of people to perform a small number of classifications each, rather than for a small number of people to each perform a very large number of classifications because, in the latter case, the data is more prone to be skewed by a single individual. This idea is summarised in the following quotation of an interview with a leading Galaxy Zoo core team member:

> **Extract 4.20:** *'There are people who have classified one million galaxies, but that's not where our data comes from, data comes from people who have done 100, and we don't get much from spurring people on to do more, the power of Galaxy Zoo comes from having lots and lots of independent people doing this. The reason why these classifications are so good is because we have 70 classifications of each galaxy' (Interview GZ-a/2)*

All of the above taken together suggests that a great deal of work and effort has been put into winning credibility for Galaxy Zoo data in the broader astronomical community. These efforts have largely been successful, as the following quotation suggests:

> **Extract 4.21:** *'[The scepticism] has decreased over the past year or two, and some people have got excited about it so people are coming and saying let's do a Galaxy Zoo with the cosmos or other surveys…it just took a few years to prove to people that this work is legitimate and very much worthwhile'* (Interview GZ-b/4)

In other words, not only has there been a broader acceptance of the Galaxy Zoo data, but the methods of Galaxy Zoo themselves have won credibility to the extent that other scientists are now seeing it as a viable way of generating data from other surveys.

The result of this is that funding bodies (such as the Leverhulme Foundation and the Science and Technology Facilities Council) have been willing to support a number of proposals from teams of astronomers to work with the Galaxy Zoo data set, and that such research has subsequently secured publication in peer-reviewed journals and has, in turn, been cited by a wide range of cosmologists.

In this Section, I have argued that members of the Galaxy Zoo core team, given their perceptions of the importance to enrol other astronomers, have been well-aware that critical tasks for them are to ensure that the project produces data unmatched by other methods and to take steps to secure its credibility in the face of scepticism in the broader cosmological community. My discussion of how they have addressed these challenges has gone some way towards explaining the success of Galaxy Zoo to date. However, the core team are not confident that future success is assured. In the next Section, I turn to contemporary developments in methods of collection and analysis of galaxy images which core team members perceive as a threat to this future success.

4.5 Looking towards the future

Driven on by the quest to understand more about the origin and evolution of the universe, and the need to base this upon robust quantitative analysis (see subsection 4.2.2), there continue to be developments in sky surveys and automated image recognition. In light of the discussion in subsection 4.4.1, these are regarded as a threat by the core team to the project's continued ability to interest astronomers in working with the project's data, in turn throwing doubt on the future of Galaxy Zoo.

This Section reviews some of these developments and argues that, understood in the context of their emerging interests with the Citizen Science Alliance, key members of the core team nevertheless have a strong motivation to secure the future of Galaxy Zoo as a flagship project of the Zooniverse. The core team are currently devoting a great deal of attention and resources towards responding to the new developments, and this bears testament to the importance to them of continuing the success of Galaxy Zoo.

As described in subsection 4.3.2, the scientific work of Galaxy Zoo is divided into two categories: the core work, which is the generation of large datasets based on volunteer classifications of galaxy images; and serendipitous discoveries, such as Hanny's Voorwerp and the Green Peas. Currently, a strong case exists that classifications by members of the public is the best available method for producing robust datasets, but rapid developments in sky surveys along with concurrent improvements in automated image recognition algorithms is considered likely by core team members to challenge this (Fortson *et al*. 2011).

One example is the Large Synoptic Survey Telescope, or LSST (LSST Science Collaboration 2009). This is currently in development, and is due to come online in the next ten years. It is anticipated that the amount of data it generates will dwarf that from other sky surveys, such as the SDSS: the LSST is expected to image approximately 50 billion objects, each object being imaged around 1000 times each, and at a higher resolution than in previous surveys (Fortson *et al*. 2011).

A set of images of this magnitude would allow for more sophisticated classifications, taking into account a larger number of criteria and providing additional statistical robustness, and any resultant dataset is therefore likely to supersede a dataset based on SDSS images. However, in order to exploit this data source fully would surely overwhelm any group of volunteers. This is explained in the following interview extract:

> **Extract 4.22:** *'The surveys we are designing now are a thousand times bigger than the Sloan so we won't have millions of galaxies, we'll have billions of galaxies. Can we use the Zooniverse technology on those? Probably not…if you want each galaxy to be observed 10 or 100 times, we're talking about 10-100 billion clicks, ain't gonna happen'* (Interview GZ-a/3)

This reflects the uncertainty amongst Galaxy Zoo team members about the project's core purpose: under such circumstances, it is feared, the project would no longer provide an unparalleled dataset for interested astronomers, and thus no longer be of interest to them[21].

[21] Indeed, many of these astronomers expressed scepticism about Galaxy Zoo's future. One, who had been a lead author on a Galaxy Zoo paper, stated that they believed the project to be *'winding down'* (Interview GZ-b/2) (**Extract 4.23**). Another said that they *'don't know'* where the project's going in the future (Interview GZ-b/1) (**Extract 4.24**).

So why might do members of the project's core team have a strong motivation to ensure Galaxy Zoo's continuation in the coming years, even in the face of new advancements in sky surveys and automated image recognition? If viewed as cosmologists only, they might, too, be expected to leave citizen science behind as it would do very little for their reputations if they were to remain closely associated with outmoded methods rather than with the cutting edge. However, they also now have another identity, namely as founders and heads of the Citizen Science Alliance, and they are now highly motivated to secure Galaxy Zoo (which they regard as the CSA's flagship project), well into the future (see subsection 4.3.3).

So how do the Galaxy Zoo core team view the task of securing the project's future in terms of ensuring that citizen science continues to be regarded as scientifically useful to cosmology? The way forward that has been proposed is to develop ways to integrate volunteer observations and automated image recognition techniques in an iterative manner (see Fortson *et al.* 2011), and this is something that is not only actively being worked upon by members of the Galaxy Zoo core team, but forms a large part of their work:

> **Extract 4.25:** *'The experiment's going to have to move from classifying everything, a brute force approach, to a smarter combination of human classification and machine classification, and how to combine those is one of the biggest research questions we're working on now'* (Interview GZ-a/1)

This might be done in two ways in particular. The first is that volunteers would classify only a subset of the LSST images: these classifications would then be used as a training set for the automated algorithms. This would be an iterative process in the sense that, first, the algorithm would be calibrated to ensure that it produced the same answers as the human observations on this initial subset of images. Then, the algorithm would run on another subset of images, and the volunteers asked to classify these images themselves to see how well the algorithm and the humans agreed on them, and so on.

The second is that volunteers will be asked to examine images to look for unusual or unexpected features (i.e. serendipitous discoveries like the Green Peas) in order to identify new classes of objects. The idea is that an algorithm could then be trained to identify instances of such a new object in order to assemble a collection of examples for further analysis (this process is described in Fortson *et al*. 2011). Even though this had not even been conceived of at the start of the Galaxy Zoo project, this is now considered a promising route to take in terms of securing citizen science involvement in cosmology, because serendipitous discoveries are only really achievable by human observation: it is not considered likely that automated methods will be developed that can identify new classes of objects, as this interview extract suggests:

> **Extract 4.26:** *'What something volunteers can always do is serendipitous discoveries...if you want to show them a bunch of pictures and say,* "Do you see something weird here?"*, there's a lot of sky to look at and this is something which probably volunteers can do much better…I don't think you can write software to pick up weird or funny things from pictures'* (Interview GZ-a/5)

Using an iterative mixture of human and automated observations, then, appears to offer Galaxy Zoo a way of continuing to produce research that is considered scientifically useful, even in the face of expected future developments, and a great deal of work to this end is being conducted by core team members. In light of these developments, the main task for these members is to ensure the project carves a unique role for itself that makes it indispensable from a scientific perspective.

4.6 Conclusion

In this Chapter, I have situated the emergence of Galaxy Zoo in light of the contexts in which its key personnel are embedded. In particular, we have seen the co-shaping of the project and the interests of these core team members and other scientists involved in the project. These interests can be seen as located in the socio-technical assemblages that constitute the project, as well as in the socio-technical linkages that constitute the broader scientific and institutional contexts in which the project is situated. How this is the case will be considered first in this concluding Section. Next, I highlight the challenges perceived by the core team to pursuing their interests. Finally, I shall look forward to the next Chapter, in which I will bring the Galaxy Zoo ethos into play and will consider the co-shaping of this ethos, the core teams' interests, and the features and policies of the project.

The co-shaping of cosmology, the Galaxy Zoo project and the interests of the projects' scientists Prior to the launch of Galaxy Zoo, the discipline of cosmology was seen as having a lowlier status compared with other branches of astronomy, in large part because it had not, to date, been possible to provide robust statistical evidence for many of the key assumptions upon which the field rests. It had been generally assumed that providing such evidence would never be feasible (and thus cosmologists had little interest in this path in any meaningful sense) until the emergence of the Sloan Digital Sky Survey (SDSS) and the development of other surveys that promised millions or even billions of galaxy images. Affording the potential of providing data to support or confirm the major hypotheses of cosmology, these surveys stimulated the interest of many within cosmology either to attempt to classify galaxy images by hand or to develop automated image recognition and classification techniques: i.e. the emergence of these surveys helped to shape and foster many cosmologists interests in providing statistical evidence for cosmology.

Aware of the limitations of these approaches for classifying galaxies, two cosmologists (Chris Lintott and Kevin Schawinski) began to conceive the Galaxy Zoo project. Prior to this, they had little apparent interest in setting-up a CCP, but the promise to these two young academics of recognition within their field that might result from producing robust datasets based on SDSS images and the example of other CCPs such as SETI@Home, stimulated an interest in running such a project. Initially, they believed that the project might attract only a small number of volunteers, primarily amateur and professional astronomers.

This expectation, however, was confounded as tens of thousands of volunteers signed up in the very early days of the project. As the project unfolded, it became apparent to core team members that they might be able to make further, previously-unanticipated, contributions to cosmology (and thus increase their prominence within the field) through the Galaxy Zoo project, by the prospects of using volunteers both to make serendipitous discoveries and to provide more detailed classifications than the core team had previously thought feasible. Thus, the core team became more interested in developing the project further to promote and encourage these, and Galaxy Zoo 2 came into being.

As a result of the increasing fame of Galaxy Zoo, both in the popular media and in the scientific world, and the growing interest in CCPs, some members of the core team started to perceive an opportunity to use their expertise (and the Galaxy Zoo platform) to help set up other projects across a range of scientific disciplines. These individuals – who had once not necessarily even been interested in running a CCP and had only initially conceived of setting up such a project in order to pursue their professional interests within cosmology – have now become interested in Citizen Cyberscience in its own right and in the prospect of leading a growing movement of such projects by setting up the Citizen Science Alliance.

Challenges past, present and future In order to pursue their various, changing interests, however, the core team members have perceived a number of challenges that they have had to address, or anticipate having to address in the near future. In particular:

- **The recruitment and retention of project volunteers** who, the core team believe, are primarily motivated by the idea that they are helping science;
- **The enrolment of other astronomers to work with the data**. The core team believe that these scientists are engaged with the project only long as it remains the best source of data to help them produce research that gains them recognition within cosmology.
- In order to successfully enrol these astronomers, **the core team have worked hard to ensure their data is viewed as credible and robust by the broader cosmological community;** and
- In order to maintain Galaxy Zoo as a flagship of the Citizen Science Alliance, and in light of perceived threats from contemporary developments in sky surveys and automated image recognition

techniques, the core team must **develop the project in such a way that it remains scientifically useful at the cutting edge of research.**

This Chapter provides a foundation for the next Chapter. In Chapter 5, I shall consider in greater depth the co-shaping of the interests of the core team the project itself, and how both of these aspects have shaped (and, in turn, been shaped by) the Galaxy Zoo ethos as the core team attempt to pursue their interests and negotiate the perceived challenges to this pursuit.

Chapter 5: The Galaxy Zoo ethos: Scientists' relationships with volunteers

5.1 Introduction

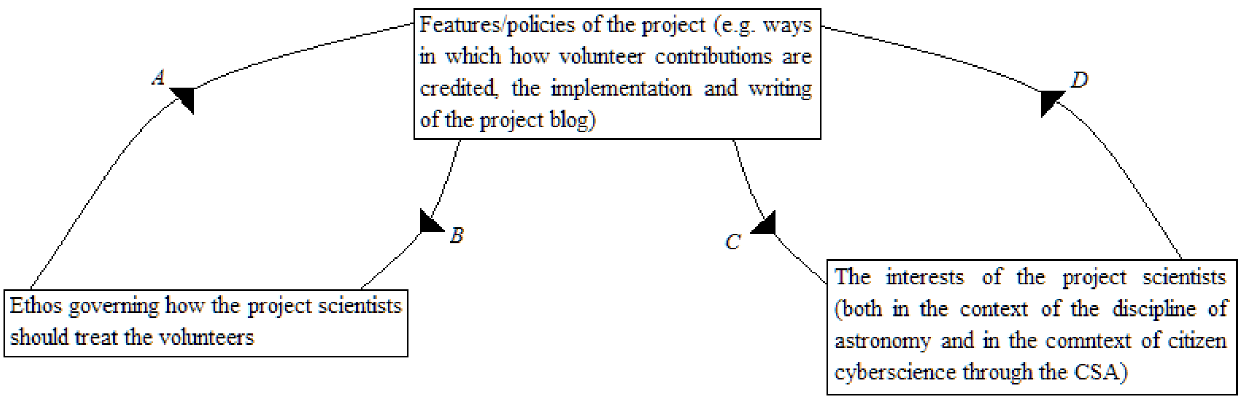
Now, I turn my attention to how the core team have sought to constitute their relationship with the volunteers, and the conditions that govern the circulation of objects and information (examples include the galaxy classifications made by volunteers and what the Galaxy Zoo scientists write on their website) amongst the volunteers and project scientists. Thus, I attempt to offer answers to Research Questions 1 and 2.

Relationships with volunteers The aim of this Chapter is to demonstrate the co-shaping of Galaxy Zoo ethos, the interests of the project scientists (see Chapter 4), and the project itself (its features, policies for how it is run etc.). My contention is that this ethos did not exist prior to, and independently of, the project, and has not been fixed over time; instead, it has emerged over the course of the project.

To see this, I will first show how this ethos is socio-technically located, in the sense that the project is constituted by the various socio-technical linkages made and sustained by the various actors involved and that the meaning of their interests is located within these links, and outside of these links, it is meaningless (this is in line with my discussion in subsection 4.1.1 in the previous Chapter about the co-shaping of technology and society). The ethos helped to guide and constrain the choices made by the Galaxy Zoo core team regarding which socio-technical linkages to make (and which to leave unmade); at the same time, the ethos has developed and become more embedded in the project as these linkages were made, in the sense that they involve ways of treating the volunteers that then become reified over time into norms of behaviour.

I will also attempt to demonstrate how the interests of the scientists are socio-technically located. Using the conclusions of the previous Chapter, I will argue that, on the one hand, the choices made regarding which socio-technical linkages to choose have also been guided by the core team's perceptions of how best to pursue their interests. On the other, the linkages that have been made have afforded the pursuit of certain pathways for the scientists to pursue some interests, and have closed off others. In other words, to reiterate the discussion in the previous Chapter (subsection 4.1.1), an individual's interests and their perceptions of how they might pursue these interests, are inextricably linked in the sense that a necessary condition for an individual to be regarded as holding a certain set of interests is that they believe that the socio-technical network in which they are embedded affords them the possibility to pursue these; on the other hand, if they do not perceive such affordances then the individual cannot be said to hold these interests in any meaningful sense. Henceforth, my use of the phrase 'an individual's interests' should be taken as shorthand for 'an individual's interests and their perceptions of how they might pursue these'. A similar approach is taken when discussing ethos, in the sense that an individual can only be regarded in a meaningful sense as having an ethos if they perceive their socio-technical network as affording them the possibility to behave in accordance with it: in the discussion below, 'an individual's ethos' should be taken to mean 'an individual's ethos and their perceptions of how they might adhere to it'.

In order to make this argument, I will first set out the key elements of the Galaxy Zoo ethos (Section 5.3). I will then examine the decision-making processes of members of the Galaxy Zoo core team in terms of how they have developed their project. At different points in the project, the core team have had to make key decisions about the features and policies of Galaxy Zoo and I will focus, in particular, on two examples of project features and policies, namely the way in which volunteers have been publicly credited for their contributions to the core scientific work of Galaxy Zoo, and the blog on the project website. Sections 5.4 and 5.5 will deal with these two examples respectively, discussing how the decision-making processes that led to the introduction and formulation of each feature and policy unfolded, showing the interplay of various factors involved (technical and scientific, ethical, and social). This is illustrated in Figure 5.1.



KEY:
A: The ethos shapes decision-making processes about features and policies of the project
B: The features and policies of the project shape the ethos, in the sense that they make it easier for, and encourage, the core team to pursue some particular approaches to treating the volunteer rather than others. This, in turn, might tend to fix these approaches as norms for the core team to behave.
C: The features and policies of the project help to shape the core team's interests, in the sense that they afford easier ways to pursue some interests ahead of others, thus strengthening the former.
D: The core team's decisions regarding which features and policies to implement, and how they are developed, are influenced by their perceptions of which choices will better help them to pursue these interests.

Figure 5.1: The relationships between the features and policies of Galaxy Zoo, the project ethos regarding how the project scientists should treat the volunteers, and the interests of the core team members.

Conditions under which objects and information circulate amongst scientists and volunteers In Section 5.6, I will to discuss how the features and policies under consideration, the norms that have developed around how they have been developed and operate by the core team (and the discourse employed by the core team when discussing this) constitutes a particular form of the *knowledge economy*, namely a *gift economy*.

5.2 The ethos of scientific practice

In order to set out a theoretical framework for the rest of this Chapter, I will first turn to the literature about ethos for scientific practice. The first thorough treatment of scientific ethos was conducted by Robert K. Merton (1949), which became the paradigmatic approach in sociology to understanding scientific practice. I will explains this work, before discussing how it was contested (and discredited) as an accurate reflection of scientific practice in later studies in the history and sociology of science (see also Chapter 2, Section 2.1).

5.2.1 Mertonian norms of scientific practice Drawing on interviews with scientists, Merton (1949) identified four norms that he argued characterised scientists' behaviour, both in public and in private scientific practice, and he argued that these norms were an objective set of universal criteria to determine whether a particular practice, theory or discipline should be considered scientific. Together, they comprise the *scientific ethos*. Subsequently, for each norm, Ian Mitroff (1974) identified a counter-norm, i.e. behaviour that was in opposition to the norm and, therefore, deemed unscientific. Merton's norms (along with their counter-norms) are:

- *Communism*, research belongs not to the individual but to the scientific community at large, and its counter-norm is *solitariness*, including the exercising of property rights and secrecy;
- *Disinterestedness,* requires that scientists seek to serve their self-interests only *'through serving the interests of a scientific community'* (Merton 1949, p. 276); this is countered by *interestedness*, namely when an individual or group seek to promote their own interests ahead of those of science;
- *Organised scepticism*, is exhibited when the scientist subjects beliefs to critical scrutiny *'against empirical and logical criteria of judgment'* (*Ibid.* p. 276); by contrast, *organised dogmatism* involves scientists asserting their own beliefs without critical self-reflection;
- *Universalism* mandates that these claims are subjected to *'pre-established, impersonal criteria'*, appealing to *'observation and previously established knowledge'* (*Ibid*., p. 276), whereas its counter-norm, *particularism*, involves judging a scientist's claim against criteria such as *'the social and psychological characteristics of the scientist'* (Mitroff 1974, p. 592).

These norms can be seen to be inter-connected. For instance, communism and disinterestedness are important for organised scepticism is to be upheld, in the sense that a scientist making a knowledge claim should be very willing to have this claim subject to intense scrutiny by others and admit the possibility they may be wrong: if they are not, it is because they are interested in achieving personal credit.

5.2.2 At what level do these norms operate? Later research cast doubt on the claim that these norms underpinned the behaviour of scientists both in public and in everyday scientific practice. For instance, in the case of disinterestedness, a number of studies have suggested that the counter-norm often drives the behaviour of scientists, putting personal goals and interests ahead of those of the scientific community. In the public arena, there have been many cases where scientists pursued individual glory or credit at the expense of other scientists (for instance, the moves by James Watson, credited as a co-discoverer of the double helical structure of DNA, to diminish the contribution of Rosalind Franklin, to this achievement (Maddox 2002)). There is also a vast literature on priority conflicts in science, where scientists engage in a struggle to secure credit as the originator of a scientific concept, or idea (see, for instance, Kragh (1980) and Prasad (2003)).

Other studies have argued that disinterestedness is violated at the very heart of everyday scientific practice. For instance, Latour & Woolgar (1979) argued that scientists' work in their laboratories is driven by a goal of securing more resources for themselves, in the face of competition from other scientists. Indeed, that scientists are able to pursue their own interests through their scientific work throws doubt on the norm of communalism: if they are able to use their own scientific output to gain additional resources, this implies some notion of propriety ownership of this output (this will be dealt with more fully in the next subsection).

Doubt has also been cast on the centrality of the other norms to scientific practice. In the case of organised scepticism, Latour (1987) suggested that the testing of knowledge claims is driven instead by social factors, arguing that scientists are involved in *'trials of strength'* with each other to reify their own knowledge claims to scientific fact (p. 78). Scientists need to enrol resources (material, financial, human) in order to contest the knowledge claims of others and to secure their own knowledge claims from the attacks of others. Those who possess these resources would only permit their enrolment if it serves their own interests to do so. What this means is that knowledge claims are contested only if sufficient resources can be enrolled to do so, i.e. the testing of knowledge claims is driven not by the norm of organised scepticism but is instead enabled or blocked by the social and political priorities of those who control the allocation of resources.

In the case of universalism, it has been argued that objective, impersonal and universal criteria for judging knowledge claims do not exist. Instead, what counts as scientific method (for instance, what counts as valid experimental procedure, what is deemed to be acceptable as evidence in scientific argument, and so on) is contingent and contested, and that experiments are not reproducible (Collins 1985).

Finally, not only have studies attempted to show that Merton's norms do not underpin scientific practice, but Gieryn (1995) argued that Merton's norms themselves are not universal and objective. Gieryn attempted to expose, instead, their historical and social contingency, arguing that they were driven by Merton's interests in attempting to discredit Freud's theory of psychoanalysis.

Thus, there is an extensive literature that suggests Merton's norms do not provide a set of objective criteria to demarcate the boundary between science and non-science. However, they are still of great interest to social and historical studies of science, because they appear to provide rhetorical tropes that are marshalled by scientists in pursuit of their own interests: they have been deployed by scientists both against the rest of society and against other scientists. This is not to say that scientists have explicitly cited Merton or specific norms by name, but that these norms are implicit in the discourse they have used.

Rhetorical deployment of Mertonian norms Gieryn (1983, 1999) argued that these norms have been used (or are implied) in the rhetorical construction of boundaries between the institutions of science and the rest of society in order to preserve and promote these institutions' claims to cultural authority as purveyors of objective truths about the natural world.

As well as being used to draw a boundary between scientific and non-scientific disciplines and practices, Merton's norms have also been implicit in rhetorical moves to construct other boundaries. For instance, they have been invoked to try and exclude certain groups of people from being regarded as scientists by playing on stereotypes to suggest that they lack the necessary cognitive abilities to think rationally, such as women (Abir-Am & Outram 1987) and ethnic minorities (Timmermans 2003).

Other examples involving science and society have been discussed in Chapter 2 (Section 2.2), where it was seen that scientists have faced objections by members of the public to knowledge claims. These objections have been dealt with by scientists attempting simultaneously to expose the interests of the lay public and the situated nature of their objections, attempting to draw a contrast with their own apparent disinterestedness, and to suggest that lay people lack the necessary training to think scientifically and thus are unable to evaluate knowledge claims objectively (i.e. they do not follow the norm of universalism).

These norms have not only been implicitly invoked in debates involving scientists and other parts of society, but also amongst scientists themselves, most notably in two types of dispute: priority disputes, and rival knowledge claims. In the case of the former, those involved have often made ad hominem attacks on rivals implying a violation of norms, for instance suggesting that a rival, driven by self-interest and a thirst for priority, was guilty of scientific malpractice by rushing to present their work through news media rather than disseminating it via the usual channel of peer-reviewed journals where its claims would be subjected to scientific scrutiny (Prasad 2007). In the case of the latter, Gilbert & Mulkay (1984) found, in a study where two groups of scientists were proponents of rival theories, that each group would describe their own work as emanating from strict adherence to the scientific method whilst casting the rival scientists' work as tainted by self-interest (violating disinterestedness, and communalism), a lack of scrutiny of their own work as they proceeded (violating organised scepticism), and sloppy experimental procedures (violating universalism).

In short, it can be seen that there is much discussion around the Mertonian scientific ethos, with many arguing that it is not a fixed and universal entity that precedes and guides all behaviour by scientists; rather the scientists should be understood as social actors, motivated at least in part by self-interest and thus whose practices frequently violate Mertonian norms. Nevertheless, as I have attempted to show above, the concept of scientific ethos is not redundant when seeking to understand scientists' behaviour, but rather that the norms can play a number of important (and context-dependent) roles in scientific practices.

5.3 The ethos of how Galaxy Zoo scientists should treat volunteers

The aim of this short Section is to set out the key norms that form the Galaxy Zoo ethos. This forms a background for the subsequent Sections, where I will attempt to chart how this ethos developed and stabilised during the course of the project. It is important to stress here that this ethos has not been fixed over time, and that, furthermore it can be regarded as indeterminate in the sense that it may be subject to future negotiation and change.

Over the past couple of years, however, there has been a great deal of consistency regarding how the Galaxy Zoo core team constitute the Galaxy Zoo ethos – i.e. the ethos has apparently stabilised. A thorough study of the blog posts, forum posts, articles and book Chapters written by Galaxy Zoo scientists, along with transcripts of interviews with these scientists, suggests that this ethos comprises four norms in particular. The first three have been well-summarised by Lintott (2010), setting out three ground-rules for Citizen Science[22]:

> *'1) Telling people about the purpose of the research and about its context is a good thing;*
> *2) Treat participants as collaborators not as subjects;*

[22] These have since been reiterated by other project scientists in talks and articles (for instance, Fortson *et al*. 2011).

3) Do not waste people's time'

The fourth is that all volunteers, and their contributions, are of *equal value to the project*.

In the rest of this Chapter, I will be dealing with three of these norms in particular. They can be seen as expressed through, and embedded in, the features and policies I consider below. One norm is that all volunteers are of equal value to the project's core scientific work, and this relates to the approaches taken in publicly crediting volunteers which treat all volunteers as equal contributors, irrespective of the number of galaxies they have classified. As will be described, the volunteers are indeed credited as equals in terms of their contributions to the core science, and no one volunteer is singled-out[23]. I then deal with two other norms, namely the first two ground-rules of Citizen Science, discussing how they relate to the project blog.

5.4 Galaxy Zoo ethos part I: Crediting the contribution of volunteers

In this Section, I focus on the decisions taken regarding how volunteer contributions to the project's core work are acknowledged. It will be seen that the core team's methods for doing this follow egalitarian approaches, in which all volunteers have an equal chance of being acknowledged. I will attempt to demonstrate the co-shaping of these methods, the Galaxy Zoo ethos, and the core teams' interests.

First, in subsection 5.4.1, I will turn to the very early days of the project when a very different approach to crediting volunteer contributions was being followed. As well as emphasizing the fact that the egalitarian approaches that have been adopted were not inevitable or fixed (and thus the core team could choose from amongst alternatives), it will also be the first step in showing the co-shaping of the Galaxy Zoo ethos, the scientists' interests and the methods for crediting volunteers, in the sense that the core team shifted from initially being unconcerned about these methods to perceiving that these methods might constrain, enable, or compromise the pursuit of interests and ethos.

In subsection 5.4.2, I will chart the further co-shaping of these methods and the ethos. In subsections 5.4.3-5.4.6, I will then argue that the core team members, when taking such decisions, also perceived an egalitarian approach to crediting volunteers as beneficial in terms of aiding the pursuit of their interests (as discussed in Chapter 4). Thus, I hope that subsections 5.4.1-5.4.6 taken together will demonstrate the co-shaping of the Galaxy Zoo core team's methods for crediting volunteers, their interests ,and the ethos.

[23] The discussion in this section is restricted to the ethos in terms of the project's core science, not serendipitous discoveries.

5.4.1 A league table of volunteers: Early implementation followed by swift abolition

At the point when the project was first launched, little thought was given by the core team regarding how to credit the contributions of the volunteers. As discussed in the previous Chapter, the core team at first anticipated a relatively small community of volunteers and had few prior assumptions about their motivations to participate, and thus few initial ideas about the policies and features of the project that might secure volunteers' continued participation and about appropriate ways to publicly acknowledge their contributions. In particular, the core team did not perceive that the question of how to credit volunteer contributions was in any sense ethically problematic.

The Galaxy Zoo project was inspired in large part by the success of many pre-existing CCPs. At the time of its launch, most successful projects were BOINC-based (e.g. SETI@Home and climateprediction.net) and it was to these that the core team turned for ideas about how to engage with the volunteers, consulting the literature produced by those involved in running these projects that detailed their experiences of Citizen Cyberscience. Much of this literature asserted that the primary reason for the continued engagement of volunteers was the existence of a system of credit, or a score assigned to each volunteer that measured their personal contribution to the data processing tasks of the project, along with league tables ranking volunteers (for instance, see Christensen *et al*. 2005). Such a system was widely- reported by scientists involved in these projects as engendering a competitive spirit amongst volunteers, encouraging them to donate more computational processing time to the project in order to improve their ranking.

As a result, the Galaxy Zoo project was launched with a league table of the top ten volunteers (ranked according to the number of galaxies classified) displayed on the front page, with the aim of providing recognition for volunteer contributions and to motivate participants. Such a system prominently acknowledged the very small number of volunteers who had clicked through the highest number of galaxies, whilst providing no recognition at all for the vast majority who could never have even hoped to classify nearly as many galaxies as those in the top 10. Indeed, as the following quotation from one of my interviewees makes clear, even those at or near the top of this league table sometimes had difficulty believing that other volunteers might be classifying the vast number of galaxies that the league table suggested:

> **Extract 5.1:** *'I remember at some point, the number one [on the league table] complained, so we had a number one guy, I made the sums, he must have been doing 8 hours a day for two months, and we had number two who was catching up and the number one said number two couldn't possibly be a real person because it's impossible to spend that much time on Galaxy Zoo.'* (Interview GZ-a/5)

After a couple of months, the decision was taken to abolish the league table, and now no trace of it exists on the Galaxy Zoo website. What happened was that the core team came to the conclusion that the league table was playing a far more important, and damaging, role than they had anticipated. This is suggested in the following extract from one of my interviews with a core team member:

> **Extract 5.2:** *'What we consciously don't want is a competition because we'd rather people classify well than classify as fast as they can. We don't want people to feel that if they classify just a few galaxies then they're not really contributing… At first, we didn't think much about putting up the leaderboard, we just thought it would be neat, but when we saw what it caused, we very quickly said that's it, it's gone'* (Interview GZ-a/2)

Here, it is suggested that the Galaxy Zoo core teams shifted from seeing methods for publicly crediting volunteers as a trivial issue to something in which their interests (in the sense that methods for crediting volunteers might impact on the quality of volunteer classifications) and their ethos regarding how to treat volunteers (in the sense that they started to see the league table as ethically problematic) are located. This is an early example of how the methods for crediting volunteers can be seen to be shaping this ethos and the core team's interests, in that through seeing the impact of how these methods operated, the core team began to regard them as important in terms of pursuing their interests and in ethical treatment of volunteers.

5.4.2 An egalitarian approach to recognising volunteers

After the abolition of the league table, volunteer contributions to the Galaxy Zoo core work started to be acknowledged in other ways, which will be discussed in this subsection. Through this, I hope to show further the co-shaping of methods for crediting volunteer contributions with the Galaxy Zoo ethos (and two norms in particular, namely that all volunteers are of equal value, and that they should be regarded as collaborators).

Egalitarian approaches to crediting volunteers As will be seen below, common to all the methods for crediting volunteer contributions that were implemented after the league table is that all volunteers are treated equally: either *all* volunteers are named in lists of contributors or, in cases where a small subset of volunteers only are named, every volunteer has an equal chance of being mentioned, however many galaxies they may have classified. This was often stated in my interviews with core team members, for instance:

> **Extract 5.3:** *'Somebody who clicked on [only] 1000 galaxies in Galaxy Zoo should feel they've contributed as much to any of the Galaxy Zoo papers as anybody else'* (Interview GZ-a/2)

The first such attempt at an egalitarian approach to crediting volunteers by name was when the first paper presenting scientific results from Galaxy Zoo was submitted to the *Monthly Notices of the Royal Astronomical Society* (MNRAS), as described in this interview extract:

> **Extract 5.4:** *'The original Galaxy Zoo paper, we tried to submit a paper with about 140 000 authors and that was rejected. This was to Monthly Notices [MNRAS]. They just disliked the fact that first forty pages of the paper were just names'* (Interview GZ-a/1)

Here, Chris Lintott included the names of all volunteers in the list of the paper's co-authors, alongside the scientists who had performed the data analysis and other members of the core team. In the event, the journal rejected the inclusion of volunteers as co-authors in this way. Instead, they were removed from the list of authors, and were listed in a random order on a page on the project's website (see Figure 5.2 which shows an extract from this page). This was produced as a poster, for which volunteers could pay a small sum to obtain, and new updated posters have been produced since, usually one per iteration of Galaxy Zoo. Although the first paper did not include a link to this webpage, it has since become standard practice to include such a link for all papers resulting from Galaxy Zoo, as is illustrated in the following extract from one of my interviews:

> **Extract 5.5:** *'We came up with the idea that any Galaxy Zoo paper could have the footnote acknowledging the users and having a link to the volunteers' page and so that's considered the extended author list, so everyone who contributed gets credit for it'* (Interview GZ-a/2)

In other instances where the volunteers are credited by name, there is space only for a limited number of volunteers to be listed. One example is when scientific results are presented as posters or presentations at conferences. In the former case, 100 volunteers are selected at random, whilst in the latter case, twelve are chosen. No volunteer is more or less likely than any other to be included, irrespective of the length of time involved in the project, the number of galaxies classified, or any other criterion. This approach has helped to reinforce the norm that all volunteers should be treated as equally valuable.

What can be seen in this (and the previous) subsection, then, is the co-shaping of the Galaxy Zoo ethos and the methods of crediting volunteers developed by the core team. In the previous subsection, I argued that very early on, the core team became aware that the league table might make volunteers feel less valuable than others. They felt that this was ethically problematic, thus developing the belief that the Galaxy Zoo ethos is located in a significant way in the methods they choose for crediting volunteers. This, in turn, contributed to their decision to abolish the league table. This is an example of how scientists' perceptions of the emergent ethos helped to shape the methods of crediting volunteers.

This initial concern with making all volunteers feel valued motivated listing all volunteers as equal contributors to the initial Galaxy Zoo paper on the website. Over time, this has become standard practice for all papers. This has helped to embed more deeply the norm that all volunteers should be considered as being of equal value, and as collaborators in the research (i.e. these methods have helped to shape the ethos). This, in turn, has promoted the introduction and establishment as standard practice of more methods which treat all volunteers as equal contributors to the project's work, and which suggest a collaborative role in some elements of this work. In other words, these are examples of how the developing ethos shaped approaches to crediting volunteers, and vice versa. As a result of these processes, elements of the ethos, namely that all volunteers are of equal value, and that they should be treated as collaborators, have become more established.

Figure 5.2: List of names of Galaxy Zoo volunteers featured on the project website
(http://zoo1.galaxyzoo.org/Volunteers.aspx (accessed 27 November 2011)).

The core teams' interests The above discussion does not fully account for the shift away from league tables and towards egalitarian methods of crediting volunteers. For instance, the many BOINC-based projects use league tables without controversy, and there is no suggestion that this is regarded as a less ethical way of crediting volunteers than the approach of Galaxy Zoo. In my study of climateprediction.net, for instance, it was very clear to me that the scientists and software engineers involved were concerned with treating the volunteers with respect, and indeed, the volunteers in that projects did not seem to feel that they were being treated badly due to the existence of the league tables. In other words, a concern on the part of a project's team to treat volunteers ethically does not necessarily lead to egalitarian approaches to crediting their contributions.

Instead, I will now argue that there is also a co-shaping of the Galaxy Zoo core teams' interests (as discussed in Chapter 4) with the methods used to acknowledge volunteer contributions. There might be seen to be a co-shaping of the various decisions made by the core team and their beliefs regarding what would prove beneficial to their interests, both in terms of establishing Galaxy Zoo as a scientifically credible project and, subsequently, in terms of addressing perceived threats to securing the future of the project. In particular, the core team believed that moving away from a league table towards egalitarian forms of credit would change volunteer behaviour in a way that would better serve the scientists' interests; conversely, (as argued above in subsection 5.4.1), the core team saw the decisions they would make about how to credit volunteers as potentially having a significant impact upon the pursuit of their interests. In particular, there are three ways in which the core team have come to perceive how their approach to crediting volunteers might be tied to the pursuit of their interests: these will be dealt with in turn in the following three subsections.

5.4.3 Improving the quality of individual classifications

As I argued in the previous Chapter, critical to Galaxy Zoo's success has been the ability to interest other astronomers in using the project data, and many resources have been expended in securing this data's credibility. The Galaxy Zoo core team believed that the league table was compromising this. As I mentioned above in subsection 5.4.1, the core team have publicly stated that a key motivation for abolishing the league table was to deter volunteers from clicking quickly through as many galaxies as possible in order to achieve a high placing on the league table. In other words, a trade-off is suggested between the speed and quality of classification, and it was believed that a league table was encouraging the former at the expense of the latter:

> **Extract 5.6:** *'What we consciously don't want is a competition because we'd rather people classify well than classify as fast as they can'.* (Interview GZ-a/2).

Another, similar motivation for removing the league table is suggested in the following extract:

> **Extract 5.7:** *'A lot of surprises came in as we progressed. The first thing is I was actually surprised by how many people wanted to cheat the system, there were just kind of trolls sending bogus clicks, writing robots just to mess up the data for pure malice in a way. Especially hackers, who just kind of what they did was they realised that if they sent this http request, they would add a click to this galaxy, so you would get hundreds of clicks for a galaxy being left-handed from the same person in like two minutes, so this obviously was wrong.'* (Interview GZ-a/5)

Here, the core team found that a number of malicious volunteers were attempting to cheat. These volunteers had become aware that they could write programmes that automatically generated large numbers of the same classification of the same galaxy, and could thus reach a high position on the league table. The impact of this was to distort the final classifications of such galaxies, in turn compromising the quality of the project data.

Hence, it became clear very early on to the core team that the league table might compromise their efforts to establish Galaxy Zoo as a credible source of scientific data, and thus threaten their ability to enrol astronomers upon whose interest the project was dependent. In this sense, then, the core team's interests in establishing the project helped to shape the methods for crediting volunteers, in the sense of encouraging them to abolish the league table.

5.4.4 Improving the robustness of Galaxy Zoo data

In addition to improving the quality of individual classifications, the core team also believed that a shift towards more egalitarian methods of crediting volunteers could assist in securing the credibility of their data in another way. Initially, as I mentioned in Chapter 4, subsection 4.3.1, the core team generally did not anticipate the number of volunteers it would attract. However, once they became aware of the scale of public interest in the project, they became interested in seeking to recruit and retain 1) a very large number of volunteers each making a small number of classifications in preference to 2) a small number of volunteers each classifying a large number of galaxies.

The reason for this preference was discussed in Chapter 4, subsection 4.4.2, namely that it helps the project to secure credibility for its data by improving its robustness. Two key assumptions of the statistical theory that has been used to process the data are that the number of classifications is very large and that each individual classification is independent of the others (and thus variations in the classifications are attributed to random statistical variation)[24]. It can thus be seen that 1) (where there are many independent volunteers making classifications) is much closer to meeting these assumptions than 2) (where the data might be skewed by systematic bias on the part of a single volunteer, e.g. if a volunteer is more likely to classify an ambiguous image as a clockwise, rather than anti-clockwise, galaxy). The following interview quotation reflects this:

> **Extract 5.8:** *'There are people who have classified one million galaxies, but that's not where our data comes from, data comes from people who have done 100, and we don't gain much from spurring a few people on to do more…The reason why these classifications are so good is because we have 70 independent classifications of every galaxy'* (Interview GZ-a/2)

The core team perceived the shift from a league table towards egalitarian methods of crediting volunteers as useful in the pursuit of objective 1), in at least two ways. One way has been discussed above, namely that abolishing the league table would slow down the rate at which some volunteers were making classifications, in order to reduce potential skew if an individual volunteer exhibited systematic bias in their classifications.

The second was a fear that a league table was fostering the impression that their contributions were not important to achieving the project's scientific goals, and this was proving off-putting to the broad mass of volunteers who contributed a relatively small number of classifications, for instance:

> **Extract 5.9:** *'We didn't want people to feel that if they classify just a few galaxies then they're not really contributing'* (Interview GZ-a/2)

[24] Many statistical theorems are generated in the limit as the size of the dataset tends to infinity. Of course, this is not a completely faithful reflection of reality, but making these assumptions is essential if data analysis is to be tractable and it offers sufficiently accurate approximations if the dataset is appropriately large.

Thus, rather than singling out a very small group of volunteers for praise, as happened with the league table, this interviewee believes that it is important to make all volunteers feel as though they are making a genuine contribution to the project science, irrespective of how many galaxies they have classified. It can also be seen here why it was felt to be important not only to abolish the league table, but also to establish other methods of crediting volunteer contributions as positive statements that the contributions really are valuable.

5.4.5 An increasing awareness of the potential for serendipitous discoveries

The two impacts anticipated by the core team of moving from a league table towards egalitarian forms of crediting volunteers that have been discussed so far can be seen as relating to attempts to establish Galaxy Zoo as a source of scientifically-credible data in the months after its launch. The third anticipated impact of having egalitarian forms of crediting volunteers discussed here is increasing the chance of volunteers making serendipitous discoveries, such as Hanny's Voorwerp or the Green Peas (see Chapter 4, subsection 4.3.2). I will argue here that such discoveries are perceived by members of the core team as not only benefitting the establishment of the credibility of Galaxy Zoo in the context of the discipline of astronomy, but also in terms of helping to secure the project's future in the context of Citizen Cyberscience and its potential use as a flagship for the Citizen Science Alliance (CSA). As I will then explain, this only became more apparent to the core team as the project unfolded. This will help to understand how egalitarian methods for crediting volunteers have become more embedded over time, as it will show how such methods has become perceived as ever more important by the core team as their awareness of future possibilities have developed.

The perceived importance by the core team of serendipitous discoveries As discussed above, the abolition of the league table was perceived by core team members as encouraging volunteers taking more time over each classification. By not encouraging volunteers to click through images as quickly as possible, the core team hope that volunteers will be more likely to make serendipitous discoveries. This is illustrated by the following quotation from one of my interviews:

> **Extract 5.10:** *'I'd rather have people…taking their time than racing through as quickly as they can, and take that extra half second for their brain to go,* "Oh, what's that blue blob down there?" *'*(Interview GZ-a/2)

So why have serendipitous discoveries been (and continue to be) regarded as important by the Galaxy Zoo core team? Two main reasons will be considered here.

Establishing the credibility of Galaxy Zoo The first reason is they have been seen by the core team as helping the Galaxy Zoo project to establish its credibility, in at least three ways. One way is that serendipitous discoveries have afforded the project's core team greater opportunities to locate the project within cosmology's mainstream. As the following interview extract explains, serendipitous discoveries have played a critical role in the development of cosmology (due to the nature of the phenomena this discipline observes) and thus the project has been able to continue this tradition:

> **Extract 5.11:** *'Some of the most interesting classes of objects [in cosmology] were initially discovered serendipitously. That's why I think serendipitous science is very important in astronomy. And the reason at the end of the day is that astronomy is a very particular science in that you can't experiment, you have one sky and you can look at the sky in several different bands. It's not like the other sciences, when you have this small lab, and you think,* "What experiment can I do to test it?" *In astronomy you can't play this game.'* (Interview GZ-a/5)

A second way in which serendipitous discoveries have helped to further the project's credibility is that the core team have sometimes been able to enrol other institutions and astronomers in helping to understand these discoveries further, thus effectively endorsing their scientific value. For instance, in the case of Hanny's Voorwerp, some institutions have been persuaded to donate resources to investigate it. For instance, a successful application was made to NASA for time on the Hubble Space Telescope (or, HST) in order to take images of the Voorwerp. The HST is vastly over-subscribed by such applications, and that such an institution approved Galaxy Zoo's application ahead of many others has been seen by core team members as a signal to the broader astronomical community that the project generates worthwhile science, as the following interview extract suggests:

> **Extract 5.12:** *'Getting this outside interest in Hanny's Voorwerp was great for us, it really meant that other astronomers were starting to take the project seriously and it got our project's name out there in the astronomy community'* (Interview GZ-a/4)

Thirdly, the core team have become aware of the potential for serendipitous discoveries in the recruitment of more project volunteers, in the sense that media coverage of, for instance Hanny's Voorwerp, is usually accompanied by a spike in volunteer recruitment:

> **Extract 5.13:** *'The media coverage [of Hanny's Voorwerp] has been really, really good. When there's a news story about the Zoo, we get an extra surge of visitors and even more statistics as people come back or come fresh to it and do more classifications'* (Interview GZ-a/7)

As discussed above, a recruitment of more volunteers is regarded by core team members as leading to a more statistically-robust data set. Indeed, the core team appear not only to be aware of the potential of Hanny's Voorwerp for volunteer recruitment, but have actually used it as a recruiting tool. For example, in August 2010, a comic book was produced under the auspices of Galaxy Zoo (Beatini *et al*. 2010, see Figure 5.3), and the inside cover of this book states, *'Discover your own Voorwerp at Galaxy Zoo'* (p. 2).

Securing the future of Galaxy Zoo In addition to serendipitous discoveries being increasingly regarded by the core team as contributing to the establishment of the project's credibility, they are also perceived as important for securing the future of the project, something in which some core team members have a growing interest in light of their setting up of the CSA. This is because such discoveries are seen as offering a possible way forward for Galaxy Zoo to continue making a contribution to cosmology in the face of developments in automated image recognition techniques for classifying galaxies (see Chapter 4, Section 4.5).

As was explained above, serendipitous discoveries have historically played a vital role in cosmology, as these have been the only way for astronomers to discover new classes of objects in space. Such discoveries can be made only through observation of galaxy images, and unlike the task of morphological galaxy classification, it is not anticipated that automated methods will be able to identify unusual images that are potential serendipitous discoveries. Thus, it is felt by some core team members that there is some potential for Galaxy Zoo to continue to pursue this route in the future, even if and when automated methods supersede the Galaxy Zoo method of classifying galaxies:

> **Extract 5.14:** *'If you want to show [the volunteers] a bunch of pictures and say,* "Do you see something weird here?"*, there's a lot of sky to look at and this is something which probably volunteers can do much better [than automated methods]. For example, for galaxy rotation and galaxy shape, you can write specialised software, but I don't think you can write software to pick up weird or funny things from pictures.'* (Interview GZ-a/5)

Although plans for the future are still at a relatively early stage, one of the possible paths being explored for the project is to assign a greater role to serendipitous discoveries: thus, fostering a culture amongst volunteers where they are encouraged to take more time over galaxy images might prove increasingly valuable to Galaxy Zoo as it seeks to entrench itself as a scientifically-worthwhile undertaking in the future.

An emergent belief in the potential benefits of serendipitous discoveries The core team regard serendipitous discoveries as having made a very important contribution to the success-to-date of the project, as well as promising further benefits in the future. However, the core team were not aware of their potential importance at the project launch; instead, this awareness only developed over time in light of the unfolding of the story of Hanny's Voorwerp, as these extracts from my interviews suggest:

> **Extract 5.15:** *'With Hanny's Voorwerp, it became apparent that serendipitous discoveries would be a big part of Galaxy Zoo. We would never have seen this thing otherwise.'* (Interview GZ-a/5)

> **Extract 5.16:** *'If you look at the original Galaxy Zoo, there was no button for* "This is weird"*, but Galaxy Zoo 2 actually has a button for* "I think this is weird" *so obviously this is something that wasn't there, at least not consciously, in the server initially. Initially, we simply wanted to use people purely as some sort of image recognition machine and I think all of us didn't quite realise the potential of what could come out of it.'* (Interview GZ-a/5)

***Figure 5.3: The front cover of* Hanny and the Mystery of the Voorwerp** *(Beatini* et al. *2010).*

What Extract 5.16 suggests is that initially, the core team only expected the volunteers to contribute classifications – indeed, the use of the word *'machine'* implies that the volunteers' work would not involve autonomy on their part, instead involving repetitive tasks according to fixed steps. When the project was first launched, there was no encouragement for volunteers to alert core team members to interesting or unusual features of galaxy images; instead, this developed over time, first by adding a short sentence at the foot of the user interface screen encouraging volunteers to share any unusual images on the online forum, and then by incorporating a button on the classification interface for the second Galaxy Zoo project. This reflects the growing awareness amongst the core team of the value of serendipitous discoveries to the project.

In short, then, the core team became aware of the value of serendipitous discoveries only some time after the start of the project, initially in terms of aiding efforts to secure scientific credibility for the project, and then in terms of helping to secure Galaxy Zoo's future over the coming years. They also link an increased likelihood of making serendipitous discoveries to encouraging volunteers to take more time and care over each galaxy classification, and perceive the shift from a league table to egalitarian methods of credit as having been helpful in this respect. Thus, while potential benefits for Galaxy Zoo afforded by serendipitous discoveries were not a driving force behind the initial move away from a league table towards more egalitarian forms of crediting volunteers, they have nevertheless provided a fresh, and growing, motivation for maintaining egalitarian approaches to crediting volunteers.

5.4.6 The co-shaping of interests and methods for crediting volunteers?

Tying together subsections 5.4.3-5.4.5, along with 5.4.1, it can be seen how the co-shaping of the interests of the Galaxy Zoo core team and the methods for crediting volunteers has unfolded. Initially, core team members did not regard these methods as relevant to their interests, but soon perceived these methods as having an impact upon the pursuit of their interests. This, in turn, gave added impetus to the move away from the league table as the core team were aware that volunteers taking more time over classifications would help to improve the quality of the Galaxy Zoo data. Then they became aware of the scale of the numbers volunteering for the project, which opened up new possibilities in terms of being able to recruit a large number of volunteers each performing a small number of classifications (rather than a few volunteers performing many classifications) which could enhance the robustness, and credibility, of Galaxy Zoo data, which provided additional reasons for choosing egalitarian methods of crediting volunteers rather than league tables. Then, they started to perceive serendipitous discoveries as potentially beneficial, first in terms of helping to establish the scientific credibility of the project and subsequently as providing a possible future for the project and, again, the core team perceive league tables as potentially harmful in this respect compared to an egalitarian approach to crediting volunteers.

In summary, this Section has discussed the co-shaping of Galaxy Zoo ethos (in particular the norms that all volunteers should be treated as being of equal value to the project, and that volunteers should be treated as collaborators and not as subjects), the interests of the project scientists (both in terms of establishing the project and ensuring its success to date, and future developments of the CSA), and the methods employed to publicly credit volunteer contributions to the project's core work. To extend this analysis, I will now turn to another feature of the project, namely its blog.

5.5 Galaxy Zoo ethos part II: The project blog

In this Section, I will attempt to demonstrate the co-shaping of Galaxy Zoo ethos (in this case, two Principles of Citizen Science, namely treating volunteers as collaborators rather than subjects, and telling volunteers about the purpose and context of the work to which they have been asked to contribute), the interests of the project's core team, and the project blog (in particular, its establishment, and the practices that have developed regarding how and when project scientists write posts for it). In subsection 5.5.2, I will argue that ethical concerns regarding how scientists should treat volunteers can be seen to have shaped decisions made by core team members about whether to implement a blog in the first place, and what should be included in the blog. Then, in subsection 5.5.3, I will argue that the blog and the blogging practices of Galaxy Zoo scientists can also be understood as being shaped by team members' perceptions of what would be most beneficial to themselves and their project. In subsection 5.5.4, I will then seek to demonstrate how, conversely, the existence of the blog and the blogging practices that have emerged also shape how the scientists perceive both their ethical obligations to blog, and how they might pursue their interests through the blog.

First, however, I will set out the key features of the Galaxy Zoo blog, including when it was implemented, who is able to contribute to it, and its content. In particular, it will be seen that the decision to implement a blog was made only some time after the launch of the project: this, along with the fact that many other CCPs (e.g. climateprediction.net) do not include a blog, should underline the point that the blog's existence (and its content) was not an inevitable feature of the project but, rather, it has emerged from a range of possibilities explored by core team members.

5.5.1 The blog's launch and its subsequent operation

The project blog was not present from the launch of the project, but instead was introduced five months later in December 2007. By December 2011, approximately 500 blogposts have been made, with over 95% being written by scientists involved in the project (of which, over half have been written by members of the core team) and the remainder contributed by volunteers. These blogposts cover a very wide range of topics concerning both the work of the project and cosmology in general.

As well as being the place where news is released about future developments of the Galaxy Zoo project and the Zooniverse (e.g. announcing the launch of new projects), the main theme of the blog is to explain the scientific progress of the project, from the initial motivation for a particular line of research all the way through to dissemination of results to the broader scientific community. For example, in the case of Hanny's Voorwerp, early blogposts focussed on initial attempts by members of the core team to make sense of the object from data that was already available (e.g. from the Sloan Digital Sky Survey), and explained why seeking further information about the Voorwerp might be of interest not only to the volunteers but to the broader astronomical community. As the research developed, posts reported back on applications to gain time on various large telescopes and at observatories in order to make observation runs and gain more data about the Voorwerp and, where successful, the blog gave accounts of these runs, even to the point of giving daily updates. The blog was also often the place to find the first release of images of the Voorwerp taken on these runs (for instance, in the case of images from the Hubble Space Telescope).

Once the scientists involved were ready to disseminate their findings about Hanny's Voorwerp to the broader community of astronomers, the blog gave details about the submission of papers about the Voorwerp. For instance, the first paper was held up by a reviewer for over a year: a blogpost gave an explanation for why this happened. Eventually, another blogpost announced its actual publication. As is standard with the publication of all papers resulting from Galaxy Zoo, one of the paper's lead authors wrote a few paragraphs attempting to explain the content of the paper in a way that is accessible to interested volunteers who may not have scientific expertise or training.

Other blogposts have been more general, giving accounts of various stages of the scientific process in general (i.e. not related to any particular paper or piece of research). For instance, there have been blogposts describing different stages of the writing of scientific papers (e.g. who actually writes the paper, who gives feedback and how this feedback is incorporated into the paper, how the peer review process operates etc.). Another series of blogposts present what is known as the *Object of the week*, in which a different interesting galaxy image posted on the forum by a volunteer is selected each week, and a more detailed explanation is given of what sort of cosmological phenomenon can be seen in the image. Furthermore, there have been a number of other blog posts on a miscellany of topics.

5.5.2 The project blog: An ethical duty to blog

Here, I will discuss how the Galaxy Zoo ethos has shaped the project blog and blogging practices of the scientists. First, I will explain that the scientists feel strongly that they have a duty to the project volunteers to blog regularly, and discuss how this relates to the ethos. Then, I will consider why they feel such a duty, before explaining how the scientists perceive their blogging practices (including what they include in the content of their blogposts) as a fulfilment of this duty. In particular, I will highlight the sometimes quite high costs to the scientists of writing blogposts, which will emphasise how seriously the scientists take this duty.

A duty to blog Blogposts are not made at the whims of the individuals concerned, but rather a culture has developed within the project where core team members and other scientists involved with the project (i.e. astronomers using Galaxy Zoo data) feel duty-bound to make regular blogposts on the topics discussed above in subsection 5.4.1. Indeed, when I was meeting with a group of interviewees who had been working with Galaxy Zoo data, they began talking with each other about how one should write a segment for the blog as they felt it had been too long since their previous post and they felt they had not been fulfilling their obligation to keep the volunteers informed of their progress. This sense of duty is further illustrated in the following extracts from my interviews with various scientists associated with the project:

> **Extract 5.17:** *'We've done the blog postings, we've been religious about that'* (Interview GZ-a/3)

And:

> **Extract 5.18:** *'We do blogposts simply because we feel obliged to'* (Interview GZ-a/6)

And:

> **Extract 5.19:** *'I was doing [the blogging] purely because I thought it was the right thing to do…not because I wanted to do it'* (Interview GZ-a/5)

I have chosen to include extracts from three different interviewees in order to demonstrate how widespread this is across the whole project. Note, in Extracts 5.17 and 5.18, the use of phrases like *'we feel obliged'* and *'we've been quite religious'* to suggest a sense of duty. Furthermore, Extracts 5.18 and 5.19 both suggest little or no personal motivation for blogging on the part of the interviewee; instead it is only really the sense of duty that compels them to do it.

So how has such a situation arisen in which the scientists feel obliged to blog about their work? Certainly, the scientists have expressed a sense of this duty from the very beginning of the blog's existence. For instance, in its very first post, the blog was introduced as a way of recognising the contribution of volunteers: because volunteers had given up their time and effort, it is only right that they are able to see and understand how their classifications have contributed to the project:

> *'Since you're the ones who have done these classifications, it's only fair that we keep you up to date on what we are doing…We created this blog to give you a window into the process by which we are conducting our research, and writing our papers'*
> (http://blogs.zooniverse.org/galaxyzoo/2007/12/24/what-this-blog-is-all-about/
> (accessed 9 December 2011))

In the aftermath of the project launch, a sense of obligation to the volunteers grew rapidly amongst the Galaxy Zoo scientists: these scientists had not initially anticipated the scale of the response from the public, but once this became apparent, they became very concerned about the possible impact of a high level of involvement in the project on the volunteers' lives:

> **Extract 5.20:** *'Early on in Zoo One [the first incarnation of Galaxy Zoo], there was a lot of concern about people spending too much time on the project and a lot of us were quite cautious…We felt we had to respect the time people were putting in...I remember once worrying that people's marriages might actually be destructed, this could be so addictive…that they spend 10 hours a day on this thing and suddenly their wife is saying,* "I'm leaving you."*'* (Interview GZ-a/3)

Humour aside, this quotation contains a serious point, namely a belief that the contribution of a volunteer represents some form of sacrifice on the part of the volunteer in the context of their own lives, i.e. they are setting aside their own self-interest to contribute to the project. This, in turn, leads to a feeling that such a sacrifice creates an obligation on the part of the project scientists to demonstrate to the volunteer that their efforts were not in vain, and the blog is regarded as a suitable medium for doing this. This is reflected in the following two quotations from the same interviewee quoted immediately above:

> **Extract 5.21:** *'We do have to write academic papers in the academic journals that speak to our colleagues in astronomy...I think that was essential to do for the volunteers so we can go back in these blog postings and we can say to our volunteers,* "Now you did make a difference, your clicks created new knowledge and it wasn't just a fun game, it wasn't just a fun exercise, it wasn't just something to keep you occupied, you together have created new knowledge that has increased our understanding of the universe"*'* (Interview GZ-a/3)

And:

> **Extract 5.22:**'*You say,* "If you give me your time, I will include you in this community and the product of you giving me your time will be a demonstration that you have changed science"… *and I think you have to complete that loop, you do have to go back through the blogs'* (Interview GZ-a/3)

What is also demonstrated by these two quotations is how the first two Principles of Citizen Science guide the blogging practices of the volunteers. The first Principle is that volunteers should be informed of the purposes of the work to which they have been asked to contribute. In Extract 5.22, the interviewee suggests that the blog is an appropriate place to do this: for instance, by blogging about scientific publications, the scientists can demonstrate that the research in which the volunteers were asked to participate has been recognised by other scientists as having a genuine scientific purpose, rather than simply being a trivial game. The second Principle is that volunteers should be treated as collaborators. In Extract 5.23, the interviewee suggests that the blogposts effectively lead to volunteers being treated as peers by expanding the boundary of who is kept in the *'loop'* about the scientific progress of the project to include volunteers: thus the blog is regarded by core team members as an identity-shaper in the sense of allowing volunteers to identify themselves as contributors to science.

The duty that Galaxy Zoo scientists feel also relates to the specific topics upon which the scientist should blog, and the manner in which blogposts should be written, as is explained in the following quotation from one of my interviews with a core team member:

> **Extract 5.23:** *'We realised very early on we should show people how scientists actually work, how science works, how we form hypotheses, how we get data, how we analyse it, how we write up a paper and the whole refereeing process, and we had a great thing saying,* "Follow the story of your work from beginning through to interpretation to publication and all that."*'* (Interview GZ-a/2)

In order to truly demonstrate to a volunteer that they have made a genuine contribution to science, the processes (and not just the output) of Galaxy Zoo's scientific work need to be opened up to them, so that they may understand how their classifications have been transformed into scientific results. In order for a volunteer to understand this, they might require a greater understanding of the scientific process (e.g. how the classifications are turned into data, how these data are analysed and turned into results by astronomers, how these results are disseminated and published to the wider community) and an understanding of the products of the process (e.g. the content of the journal articles that have been published, and how they contribute to the advance of astronomical knowledge in general).

The cost to Galaxy Zoo scientists of blogging The obligation that scientists feel to the volunteers is, however, reflected not only to the existence of the blogpost or in the topics covered, but also by recognising that the production of this blog is regarded by the scientists as costly and potentially risky to their interests. In other words, the sacrifice of the volunteers is acknowledged and reciprocated by a genuine sacrifice on the part of the scientists in writing blogposts. In order to illustrate this, two costs will be explained here. The first is that writing the blogposts is regarded as hard work by the project scientists. For instance, when presenting a paper they have written in the blog, a scientist is expected not only to provide a summary of its contents but also to explain often very complicated and esoteric concepts in ways that might be grasped by volunteers with little or no prior scientific knowledge and training. This is considered in the following interview extract:

> **Extract 5.24:** *'I was doing [blogposts]…not because I wanted to do it, and I find presenting our results quite hard. What we measured was whether, if you find two galaxies close together, that they're more likely to spin the same way, and the mathematical theory behind it is so dense… I just tried to convey the science we did in as simple a manner as possible'* (Interview GZ-a/5)

Here, the interviewee talks about how difficult it is to construct a blogpost that is easily comprehensible to volunteers. What is interesting to note is that the interviewee does not appear to enjoy writing such posts; rather she appears to feel obliged to do so. In particular, the interviewee does not believe that she can get away with a post that might be beyond the grasp of the volunteers – instead, she must work hard so that the blogpost is as comprehensible as possible. In other words, the interviewee is suggesting a duty that, in writing a blogpost, she should not simply be going through the motions for the sake of appearing as though she is feeding back to the volunteers, but rather that a blogpost should involve a true giving back of something that is of genuine value to volunteers (i.e. something that they can truly appreciate).

The other cost that the scientists have faced as a result of the blog is that, by being so open about their results (including unpublished results), they run the risk of other scientists from outside the project using ideas on the blog in their own work, raising the possibility that Galaxy Zoo scientists might find that other scientists beat them to publication of results. Indeed, in the event, this has happened on at least one occasion, in the case of Hanny's Voorwerp (see subsection 5.4.4). However, the core team felt that this was a risk they were willing to take, as this interview extract suggests:

> **Extract 5.25:** *'Astronomy is very competitive and anybody could look up on the blog and write their own paper if they take our ideas off the blog, but it's a risk we decided to take because we thought this is the citizen scientists' discovery as well…I felt we owed them this channel of communication'* (Interview GZ-a/2)

By asserting that the Galaxy Zoo scientists are willing to take such a risk in order to write blogposts, the interviewee is suggesting how seriously they take their obligation to recognise and reciprocate the volunteers' contributions to the project. In other words, the scientists' obligation to blog trumps the scientists' pursuit of personal prestige within their academic community.

However, although the scientists have justified the blog's existence in terms of ethical purposes, this does not provide a full account of how it came into being or the development of blogging practices by the project scientists. After all, few of the BOINC-based CCPs have a similar type of blog giving details about the processes by which volunteer contributions are turned into results. Sure enough, these projects have pages on their website which give an account of the scientific aims of the project, and some of the science underpinning the project, but these tend to be static over time, rather than detailing specific papers and attempting to explain them in terms accessible to laypeople. These projects' websites also have news pages giving details of new developments in how the project is run, and announcing the publication of scientific articles (and sometimes scientists will post on a project's forum with similar news) but the frequency of news stories does not approach the level of activity on the Galaxy Zoo blog. Thus, it can be seen that a general concern with treating the volunteers with respect does not necessarily lead to the establishment of a blog, nor to the content regarding the opening up of the scientific process that is seen on the Galaxy Zoo blog.

5.5.3 Establishing and writing the project blog: Pragmatic moves

Instead, it can be seen that the blog's development is linked not only with the development of the ethos, but also with the emerging interests of Galaxy Zoo scientists as the project unfolded. As I have explained above, the scale of the public response to the project took core team members by surprise, which made them aware of the possibility of recruiting a large corpus of volunteers. This promised to improve the robustness of the project's dataset, in turn promoting the success of the project by helping to secure the data's credibility and thus interesting teams of astronomers to use the data. In my interviews, the project scientists consistently cited the blog as a key way of keeping volunteers engaged, as exemplified in the following interview extract[25]:

> **Extract 5.26:** *'I felt that communicating back with them helps drive the project forward and gets people more interested in the project'* (Interview GZ-b/2)

Since the early days of the project, many members of the core team developed a belief that the project volunteers were motivated to participate by the idea that they would make a real contribution to this science, i.e. by allowing the volunteers to identify themselves as contributors to science. This early belief was fostered when, within two weeks of the project launch, Kate Land (a member of the Galaxy Zoo core team) started up a subforum in which volunteers were asked to respond to the following:

> *'We feel like we've really captured people's imaginations with Galaxy Zoo, and we'd like to know why. Please leave us feedback here about why you are taking part.'* (http://www.galaxyzooforum.org/index.php?topic=68.0 (accessed 9 December 2011))

[25] In the previous Chapter (subsection 4.4.1), I explained how some scientists involved with Galaxy Zoo have expressed scepticism that the blog is a way of engaging with a large number of volunteers. Just to clarify, there I was talking about those astronomers from outside the core team who work with the project data. Here, I am focussing instead on core team members, as they are the ones who took the decision to implement the blog and provided most of the early (and many of the subsequent) blogposts in which the norms and culture regarding blogging were established – as I mentioned in subsection 4.4.1, these individuals generally appear much more enthusiastic about the blog's potential for engaging with a large number of volunteers.

Many volunteers have since responded, and approximately half who stated explicitly that they were motivated to participate by the possibility of contributing to the advance of scientific knowledge[26]. This is exemplified in the following post made by a volunteer in reply to the post by Kate Land:

> *'The opportunity to work with professional astronomers and help out with real research is the nearest I will ever get to "being" an astronomer and contributing to new discoveries'*

That this is the primary motivation for volunteers has since been further reinforced in the minds of the core team by surveys of the volunteers. For instance, in January 2008, some members of the core team started to work with social scientists in order to understand the motivations of volunteers, and this offered support to the idea that it was the notion that they would contribute to the scientific work:

> **Extract 5.27:** *'We have some social scientists working with us and the number one question we wanted to know the answer to was,* "Why do people want to do this, and why are they so excited and who are these people?"*, and so the upshot of all this analysis is that ... the number one motivation is because people want to contribute to real science. They didn't want to go to some public outreach website and learn about the work of scientists, they didn't want the finished product so to speak, but to take part in the enterprise of science'* (Interview GZ-a/2)

Understood in these terms, the blog can be seen not only as a way of demonstrating to the volunteers not only the publishable results arising from the project, but showing how their classifications contribute to journal articles and conference presentations by explaining the day-to-day processes which lead to these finished products and the role of volunteer classifications in this. The process of getting from the generation of data to the publication of results is often subject to numerous delays for a variety of reasons, and giving details of the process along the way allows for more regular feedback to the volunteers which is seen as helping to reassure them that scientific progress is being made (and thus to keep them engaged) even if no papers have yet been published, as the following interview extract with a core team member makes clear:

> **Extract 5.28:** *'[The blog] is also to keep people on board, to keep them interested and coming back…We've been using all their data, showing how all the clicks they've done are being used in scientific terms because the gap between collecting data and getting something useful from it [is long]…and we wanted to show them how useful the information was'* (Interview GZ-a/6)

As well as helping to create an impression of progress by the project even in the absence of publications, the blog has also helped to reassure volunteers about such an absence by offering explanations for delays and suggesting they are a normal – even desirable – part of science. For instance, in the case of Hanny's Voorwerp, the paper was held up in the peer reviewing process by nearly a year (see subsection 5.4.4). In a blogpost, Chris Lintott explained away the delay thus:

> *'The fifth Galaxy Zoo paper – the one that discusses Hanny's Voorwerp – has now been accepted for publication by Monthly Notices of the Royal Astronomical Society. It's somewhat of a relief to say that, as it was way back in August of last year that we first submitted it. The referee was extremely thorough,*

[26] Other common motivations included the opportunity to learn more about astronomy, an enjoyment of the galaxy images, and seeking to understand better the place of humans in the universe.

catching a few stupid mistakes we'd made (as a good referee should)' (blogs.zooniverse.org/galaxyzoo/2009/06/29/voorwerps-everywhere (accessed 9 December 2011))

He also linked to another blogpost which gave a general account of the peer review process. As a result, a volunteer interested in the progress of Hanny's Voorwerp, and who might be wondering whether any progress had indeed been made (and thus might be questioning the scientific value of the project), could get reassurance that the delay is instead an integral part of the scientific process and is, indeed, making the scientific output of the project even more robust and scientifically valuable.

5.5.4 The shaping of the Galaxy Zoo ethos and the core team's interests by the blog

In the introductory part of this Section, I claimed that the Galaxy Zoo ethos, the core team's interests, and the blog and blogging practices were all co-shaping each other. So far, I have attempted to show that the ethos and the core team's interests have shaped the blog, but not the other way round. Sure enough, it would seem (from my interviews) that both a sense of obligation to feed scientific progress back to the volunteers and a concern with keeping volunteers engaged in order to help advance the project's scientific goals preceded and motivated the establishment of the blog. Nevertheless, it would appear that the existence of the blog and the practices associated with it have also shaped the ethos and the interests of the core team by helping to entrench the former and have reinforced the idea amongst the core team that the blog is a good way to keep volunteers engaged. This is expressed in the following extract from an interview with a core team member:

> **Extract 5.29:** *'Now we really have to keep updating the blog. The volunteers ask us on the forums when they haven't heard anything for a while, now they expect us to update regularly the blog so it seems rude if we've been quiet…We also don't want them getting suspicious if they haven't heard anything, wondering if we've stopped doing the science'* (Interview GZ-a/7)

Here, then, the core team member worries that the blog's existence, and its regular updating, has created an expectation on the part of the volunteers that the scientists should update the blog: to do otherwise, this interviewee suggests, risks both offending the volunteers and jeopardizing their continued involvement in the project by making them wonder whether their contributions are continuing to lead to scientific results. This, in turn, entrenches both the notion that it is respectful to the volunteers to make regular blogposts (i.e. it reinforces the form of ethos discussed above in subsection 5.5.2) and the idea that regular blogposts are required to keep volunteers engaged (i.e. to help in the pursuit of the core team's interests).

5.5.5 Exchanges between the Galaxy Zoo scientists and the volunteers

In this Section, we have seen ways in which the co-shaping has occurred of the Galaxy Zoo ethos (and two of the Principles of Citizen Science in particular), the interests of the project's core team, and the blog and blogging practices of these scientists. This, in tandem with the arguments of the previous Section, demonstrate how both the ethos and the interests of the Galaxy Zoo core team have not been fixed since the launch of the project (and thus cannot be seen simply as causal factors in the development of the project), but instead can be regarded as being located in complex socio-technical networks that have emerged and stabilised as the project has unfolded.

At the heart of these networks are exchanges between volunteers and scientists, and the conditions which regulate the terms of these exchanges. In this and the previous Section, we have seen how volunteers give classifications of galaxy images to the scientists and in return, the scientists give blogposts and public recognition to the volunteers. The conditions of exchange of these objects have developed and become stabilised during the course of the project, and they are located in the methods of crediting volunteers, the project blog and blogging practices, and the discourses of the project scientists surrounding these.

5.6 The *'knowledge economy'*[27] of Galaxy Zoo

The phrase 'knowledge economy' can have a number of meanings, depending upon the context in which it is used. In this case, it will refer to the circumstances and ways in which knowledge is exchanged between individuals and institutions, focussing on the case of scientific knowledge. It refers to the conditions of exchange, for instance, notions of ownership of, and responsibility for, knowledge claims. It is also concerned with the notion of value of knowledge claims: in what sense can a claim be said to have a value, and for whom, and how does this value emerge? In the case here, scientific knowledge does not just refer to that which is circulated in formal publications (e.g. academic journals or papers presented at conferences) but throughout various stages of knowledge production, for example the transfer of data between people in a collaboration, ideas exchanged in informal conversations amongst colleagues, or methods of crediting the contributions of others (for instance, through co-authorship in a journal article).

McSherry (2003) has identified in the academic literature two models for scientific knowledge economies. One is the *gift economy*, and the other is a system of capital accumulation, perhaps best conceptualised as a *market economy*. Each of these terms will be explained and discussed in turn below. Further, each model will be linked to the discussion of scientific practice in the previous subsection: it will be seen that the behaviour and values

[27] McSherry (2003, p. 225)

displayed by individuals involved in a hypothetical gift economy correspond closely to the Mertonian norms of scientific behaviour (see Section 5.2 above), whilst those in a market economy of scientific knowledge implicate some of the counter-norms. Finally, I will discuss how a gift economy of knowledge involving both project scientists and volunteers has emerged in the project.

5.6.1 The gift economy of knowledge

The gift form of knowledge economy is founded on the idea that an individual or institution cannot exercise property rights over a piece of scientific knowledge, and that to seek to do so is ethically wrong; instead, scientists have an obligation to share scientific knowledge with other scientists, to ensure that their knowledge claims are accurate and not misleading, and reciprocity in the sense of a scientist acknowledging what has been given to them by others (Hagstrom 1965). This is premised on the notion that it is absurd for scientific knowledge to belong to anybody, in the sense that a scientific knowledge claim is merely a description of natural phenomena that exist independently of, and prior to, a claim. To highlight this notion, Biagioli (1998) contrasts the scientific author with the author of a work of fiction: whilst the latter is able to assert property rights over their work because the work is a product of the author's creativity, a scientist asserting similar rights would be suggesting that their work was a product of creativity, which would throw into doubt the objectivity and universal truth value of this work. That a scientist should not exercise property rights over a knowledge claim means they should share this knowledge freely with other scientists.

Thus, in a gift economy, a scientist's name being attached to a knowledge claim does not signify ownership. Instead, it functions as a marker that this knowledge claim has been made in good faith and has been arrived at, not through fraudulent practices, but by sincere work on the part of the scientist making the claim. It is the scientist's reputation that guarantees the veracity of this knowledge claim (and deters fraud): should this claim be exposed as fraudulent, their reputation would be destroyed and they would face exclusion from the scientific community (Biagioli 2003).

A final feature of this gift economy is what it means to say the gift has value. The value is not understood in terms of its use to the recipient or to the broader scientific community; rather, it is endowed with value by the donor. The donor is assumed to have invested time and effort into producing it, ensuring its veracity to the best of their ability and means and investing their reputation in this, and releasing it without expectation of credit in return is seen as a sacrifice (McSherry 2003). This, in turn, mandates acknowledgment and reciprocity on the part of the recipient: to do otherwise would be seen as impolite.

It can be seen that the Mertonian norms described in subsection 5.2.1 are integral parts of the gift economy of knowledge. For instance, communalism is the notion that scientific knowledge does not belong to an individual or institution, but to the entire scientific community. This leads to disinterestedness, in the sense that scientists are not able to use the scientific knowledge that they generate to promote their own interests, motivated instead by advancing scientific knowledge. The norms of universalism and organised scepticism manifest themselves in the sense that, because the name of a scientific author acts as a guarantor of the knowledge claims with which it is linked, the author should be trusted to have subjected their claims to extensive scrutiny, using well-established methods of scientific inquiry to ensure the veracity of their claims.

5.6.2 Market model of the knowledge economy An alternative model of the way in which scientific knowledge is exchanged is the market model. Bourdieu (1988) discussed the notion of *'symbolic credit'* (p. 284), in which knowledge claims could be used by those making them to accrue credit and prestige, suggesting that academics are motivated by accumulating more symbolic capital for themselves and pursue this through their work.

Latour & Woolgar (1979) adopted a similar theme in their study of scientists in a laboratory. They suggested that scientists are motivated by the pursuit of greater recognition. They are engaged in *'cycles of credit'* (p. 187), in which their work in the laboratory has as its goal the publication of articles, leading to increased recognition as credible practitioners of their field, which, in turn, leads to a greater ability to procure financial and material resources to pursue more scientific work (see Figure 5.4). In such an knowledge economy, scientists can be seen to exercise some form of property right over their work and the knowledge claims they produce, in the sense that by attaching their name to a piece of work, they are able to trade on this to accumulate more resources for themselves.

Credit in such a market knowledge economy serves a very different function than in a gift economy. In the former, credit serves as a something that a scientist can invest and transform into money/equipment etc. Here, the value of a scientist's knowledge claim is determined by the amount of resources that it allows the scientist to subsequently procure, i.e. by extrinsic factors. By contrast, the function of credit assigned to a scientist in a gift economy has the function of honouring and recognising the scientist's work and personal sacrifice involved, i.e. its value is intrinsic to the scientist producing the knowledge claim.

In contrast to the gift economy, a market economy can be seen as violating a number of Mertonian norms, implicating their counter-norms, in particular solitariness (where property rights over pieces of knowledge are assigned to, and exercised by, groups and individuals) and interestedness (where groups and individuals act in their own self-interest, ahead of the interest of the collective-as-a-whole).

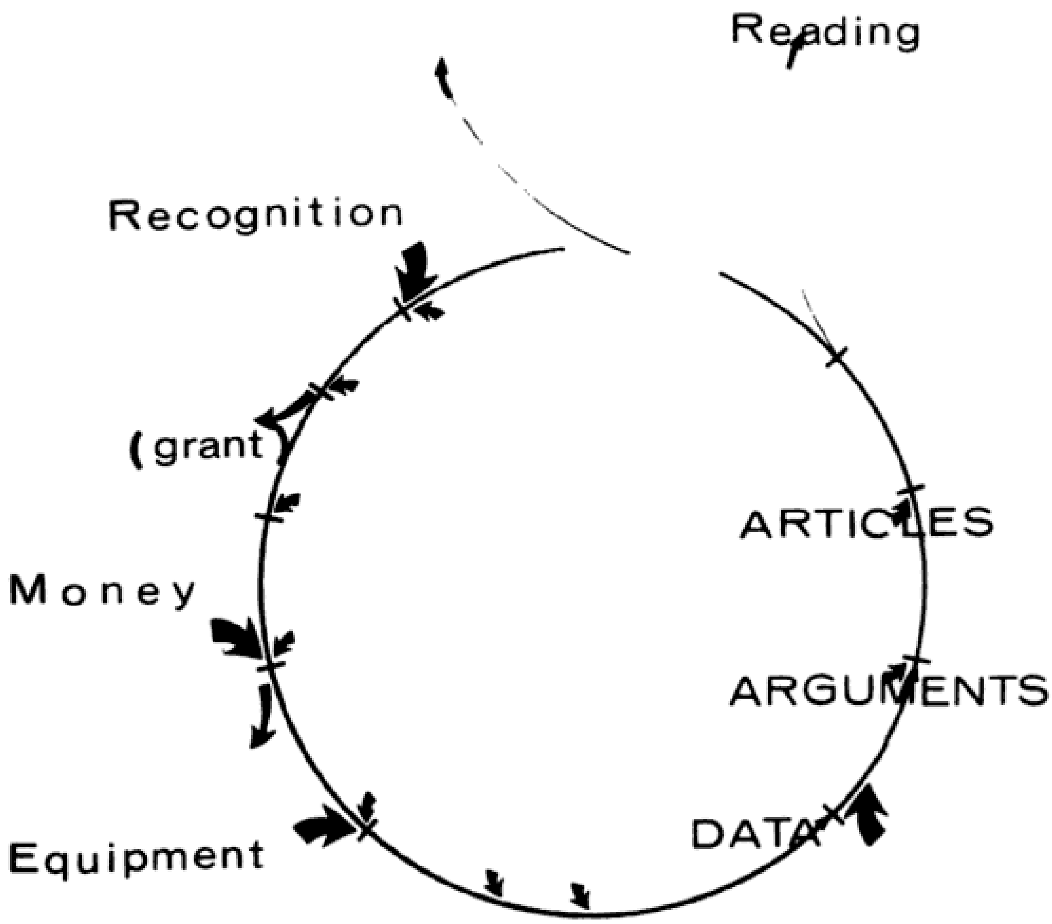


Figure 5.4: Cycles of credit. *Going clockwise from the bottom right-hand corner, this diagram illustrates how scientists use the data they produce to construct scientific arguments, which are then presented in published articles. Such publications in turn lead to greater recognition in the broader academic community, thus improving the scientist's chances of obtaining further funding to buy new equipment/infrastructure, with which to produce more data. Thus the cycle begins again (Image taken from Latour & Woolgar (1979, p. 201)).*

5.6.3 The knowledge economy and forms of crediting volunteers in Galaxy Zoo

The shift away from a league table towards egalitarian approaches to crediting volunteers (see Section 5.4 above) has contributed to the embedding of a gift economy of knowledge for the circulation of objects and knowledge amongst volunteers and Galaxy Zoo scientists in a number of ways.

One way is that egalitarian approaches suggest that the value of a volunteer's contribution should be understood in terms of the sacrifice made by the volunteer, rather than being assessed in terms of its value to the recipient. To see this, consider how the league table, by contrast, violated this key feature of a gift economy and instead pointed towards a market economy. When Galaxy Zoo assigned a numerical value to the contribution of a volunteer (number of galaxies classified), this signalled that the value of a donor's (a volunteer's) contribution was determined by its value to the recipient (the project), rather than in terms of what it meant to the volunteer.

Furthermore, not only did a shift from league table to egalitarian forms of credit remove this violation of a gift economy, it also actively helped to constitute a gift economy by suggesting that the work of every volunteer is valuable: after all, the core team could have replaced the league table with nothing that credits volunteers by name, which would leave open the possibility that a volunteer's contribution might be worthless. By naming every volunteer, the core team suggests that the classifications of each volunteer represent a sincere and valuable donation, irrespective of how few or inaccurate, they might be.

The discourse of the core team when discussing the abolition of the league table to volunteers and in other public forums has also assisted in constituting a gift economy. Galaxy Zoo scientists have, in public, cited one motivation above all others for moving from the towards more egalitarian forms of crediting volunteers, namely that giving greater prominence to those who had classified more galaxies would act as an incentive for some volunteers simply to click through as many galaxies as quickly as possible, thereby compromising the quality of classifications (and the data produced by the project):

> **Extract 5.30:** *'We'd rather people classify well than classify as fast as they can, we tell them many times on the forum,* "This is not a competition and we'd rather you'd spend the time doing it as well as you can rather than race towards a goal." *Very early on we had a leaderboard…this turned out to be a really, really bad idea…we don't want it to be a race'* (Interview GZ-a/2)

Although this reason is cast in terms of the scientific credibility of the project (and thus might be understood in terms of the scientists' self-interest), it also serves to reinforce the notion of a gift economy because it suggests that it is precisely the violation of such an economy (in favour of a market economy) that compromises the science: in other words, a league table was cast as discouraging the Mertonian norm of disinterestedness[28], which could potentially lead to poorer quality data, in the sense that it was encouraging volunteers to rush and thus take less care in making classifications (i.e. by not taking the time to subject each galaxy image neither to a proper degree of scrutiny, thereby contravening the norm of organised scepticism, nor according to the appropriate scientific criteria, thereby contravening universalism).

[28] In subsection 5.2.2, I discussed how the norm of disinterestedness is implicit in a gift economy.

Finally, the way in which the egalitarian forms of naming volunteers has suggested that volunteers are, in some sense, scientific peers (for instance, by making them visible by name to other members of the astronomical community, or implying they are co-authors on papers) has reinforced the notion that they are included in the same knowledge-making community (and thus, knowledge economy) as the scientists. As I discussed in subsection 5.2.2, the gift economy is held to be the conditions under which knowledge (and tools for generating this knowledge) circulates amongst the scientific community and thus suggesting that volunteers are scientific peers in turn suggests this gift economy has expanded to include these volunteers.

These are not the only ways in which a gift economy of knowledge has become embedded in Galaxy Zoo: in the next subsection I will explain how the blog and associated blogging practices also accomplish this.

5.6.4 The knowledge economy and the Galaxy Zoo blog

Here, I will highlight how this constitutes two key features of a gift economy. The first is that an item being exchanged is endowed with value because it involves a sacrifice on the part of the donor, and it is this value that mandates something being given in return that involves a sacrifice on the part of the recipient. As I explained in subsection 5.5.2, it can be seen that the scientists have come to regard the volunteer classifications as mandating a recognition through blogposts. In particular, writing blogposts can be seen as a sacrifice by the scientists both in terms of the work involved and the risks involved in openly discussing current lines of research.

The second is that the blog can be seen as a sharing of scientific knowledge with the volunteers, suggesting a common ownership of this knowledge by both the scientists and the volunteers, in the sense that not only do the volunteers get to see the finished results that are disseminated to the broader community but are also party to the interim stages of research (and, indeed, could give feedback through the comments section on the blog). In particular, the two costs to the scientists of writing the blog discussed in subsection 5.5.2 can both be seen as helping to constitute this element of a gift economy of knowledge in different ways. The first cost, namely the often painstaking effort involved in trying to ensure that the science is made fully comprehensible to volunteers, can be understood in this way, in the sense that unless a volunteer can truly grasp the science of a blogpost then they cannot be truly be said to have ownership of it in a cognitive sense.

The second cost is the risk of being scooped by scientists from outside the project as a result of making public interim stages of research on the blog. The scientists' willingness to bear this cost can also be seen to constitute a gift economy, in the sense that they would rather share their work with the volunteers (and thus reaffirm the notion that the project's science is held in common by both scientist and volunteer alike) than keep information about their work to themselves in order to avoid being scooped and gain greater personal recognition (which would effectively involve the assertion of individual property rights over the science on the part of the scientists, as in a market economy).

5.6.5 Galaxy Zoo and Mertonian norms of scientific ethos

In this Section, I have attempted to demonstrate how the socio-technical networks of Galaxy Zoo (and in particular those which have emerged during the co-shaping of the blog and methods for crediting volunteers, the Galaxy Zoo ethos, and the interests of the project scientists) constitute a gift economy of knowledge for the circulation of objects and information amongst volunteers and project scientists.

As I discussed above in subsection 5.6.1, this form of the knowledge economy is inextricably linked with Mertonian norms of scientific ethos. If exchanges amongst a group of individuals operate in accordance with the conditions of a gift economy, then the group is adhering to these norms; conversely, if a group is concerned with adherence to such norms, then their behaviour when making exchanges amongst themselves will call into being a gift economy. It might seem, therefore, that not only has the Galaxy Zoo ethos emerged and been stabilised, but that Mertonian norms have become embedded in the project.

However, it would be wrong to conclude simply that the Galaxy Zoo scientists do indeed operate by Mertonian norms in a non-problematic way when dealing with volunteers. It should be noted that a key idea underpinning these norms (and the gift economy of knowledge) is that scientists are willing to put their own self-interest (and ideas of glory and self-promotion) to one side and instead become passive instruments, serving the interests of science and the scientific endeavour. What we have seen here instead is perhaps ironic: whilst Mertonian norms and the gift economy of knowledge have appeared to emerge and become stabilised in the scientists' dealings with the volunteers, this is only as a result of choices made by the core team members in accordance with their perceptions of what would be in their own interests.

As stated in Section 5.1, the Galaxy Zoo ethos should be regarded as indeterminate, in the sense that it is subject to a continuous process of negotiation and possible change. Certainly, it appears to have stabilised so far, but as we have seen, the key decision-makers in Galaxy Zoo are aware of their own, changing personal interests and how they might best pursue these. Thus, it is possible that were they to believe that changing the way in which they interact with volunteers might better enable them to pursue their interests, then the ethos might change.

5.7 Conclusion

Drawing on the conclusions of the previous Chapter, I have sought to answer RQs 1 and 2, considering how, in the context of Galaxy Zoo, the project core team have engaged with volunteers. I have argued that a particular ethos regarding how project scientists should treat volunteers has emerged and become stabilised, and that a gift form of the knowledge economy governing exchanges between volunteers has become established.

Processes of co-shaping Through particular consideration of two features of the project (the blog, and approaches for acknowledging volunteer contributions), we have seen the co-shaping as the project unfolded of these features of the project, the Galaxy Zoo ethos, and the interests of the project core team (see Figure 5.1).

This Chapter documented how a general concern on the part of the core team for respecting the volunteers was a factor in driving the establishment of the blog and the development of specific blogging practices (recognising the sacrifice of the volunteers), as well as being a major factor in the decision to move from a league table crediting only a handful of volunteers to an approach that presented all volunteers' contributions as equal (so that no volunteer felt unvalued). At the same time, the development of these features and practices helped to shape, refine and reinforce the ethos, taking it from a general concern with respecting volunteers to specific, codified norms regarding the treatment of volunteers.

However, by considering other CCPs, it also became clear that simply considering the co-shaping of the Galaxy Zoo ethos and the project's features and practices was not enough to explain what had happened in Galaxy Zoo Drawing on conclusions for Chapter 4, I argued that there has also been co-shaping of the interests of the Galaxy Zoo core team and the features and practices of the project.

In particular, it can be seen that the interests of the core team, and their perceptions of how they might pursue these and what challenges to this might arise in the future from automated image recognition techniques, have changed as the project unfolded (e.g. shifting from being interested in their reputations in cosmology to leading a movement of Citizen Cyberscience). Over time, the core team have become more convinced of the value of blogging and egalitarian methods (rather than league tables) for crediting volunteers to the pursuit of these interests, and thus these approaches have become more entrenched. Conversely, I argued that the way in which these approaches have developed have, in turn, shaped the interests of members of the core team, for instance they initially regarded methods of crediting volunteers as a trivial issue before coming to the view that these methods aided the pursuit of particular interests (e.g. volunteer recruitment and retention, or slowing classifications down).

The Galaxy Zoo core team What can be seen in this Chapter, then, is an issue relating to the behaviour of the Galaxy Zoo core team. Unlike the inventor of the Zimbabwean Bush Pump (q.v. Chapter 4, subsection 4.1.1), the core team members do seek success (in the contexts of cosmology, Citizen Cyberscience etc.) through interesting and enrolling others into their networks, and marshalling these networks: they are not content to simply *'dissolve'* into the background (de Laet & Mol 2000, p. 227). At the same time, however, the core team appear to be different from the actor of Actor-Network Theory, who is concerned only with pursuing their own self-interest and for whom allies can be discarded as soon as they are no longer regarded as useful in this pursuit (Latour 1988). What we have, in the case of the Galaxy Zoo core team, are a group of individuals who are both ambitious for themselves and for the institutions to which they belong, but also have a sense of obligation to the volunteers: their explorations of the various possibilities afforded to them through the unfolding project and the decisions they have had to make about how to develop the project have, at their heart, involved a drive to ensure simultaneously that the core team can pursue their interests and that they treat the volunteers with respect.

Gift economy of knowledge and norms In this Chapter, I also discussed how the blogging practices, the approaches to crediting volunteers, and the discourse of the core team surrounding these, have come to constitute a gift economy of knowledge governing exchanges amongst project scientists and volunteers. In particular, it can be seen that such a knowledge economy is not an essential feature of a CCP, but has instead only come into being through the particular socio-technical configurations that have developed.

Through this can be seen the way in which Mertonian norms have come into being, and have become embedded in the ways in which the project scientists treat volunteers. As discussed in subsection 5.2.1, Merton (1949) argued that his norms of scientific practice were abstract entities that exist independently of, and prior to, scientific practice and act as constraints on scientists' behaviour that would completely override any pursuit of self-interest: this is not the case here. Nor, however, is it the case in Galaxy Zoo that these norms have only been deployed, or implied, at a rhetorical level by the core team, as might be expected from literature discussed in subsection 5.2.2 (Gieryn 1999). Instead, what has happened has been more complex: sure enough, these norms appear to have emerged and stabilised (rather than being fixed over time, and prior to the project) as the project has unfolded and (in particular) as approaches to crediting volunteers and to blogging have developed over time. At the same time, however, they have come about in a non-Mertonian way, i.e. as a result of choices that have been made by the core team in accordance with what they believed to be their self-interest. These norms then, are not an inevitable feature of the project, nor will they necessarily bind Galaxy Zoo scientists for the rest of the project's life as these scientists' interests (and perceptions of how to pursue these) change. The Galaxy Zoo ethos, instead, is indeterminate and subject to constant negotiation and possible change. That this is the case will be emphasized in the next Chapter through consideration of another CCP, climateprediction.net, where a different approaches regarding how project scientists should treat volunteers have emerged.

Chapter 6: climateprediction.net and approaches to volunteer retention

6.1 Introduction

The following two Chapters present some findings from my case study of climateprediction.net. I seek to examine how the credit and other statistics systems operate in climateprediction.net, and will set out to show that they have played a critical role in the project as experienced by a variety of project personnel (both volunteers and project team members, i.e. the scientists and software engineers involved in setting up and running the project). As in the case of Galaxy Zoo, climateprediction.net was launched with a league table of volunteers, measuring the amount of work they contributed to the project. In the case of climateprediction.net, however, the ranking was not based on the number of galaxies classified but on the quantity of computational processing of climate models performed on a volunteer's computer (measured in *credits*). Unlike Galaxy Zoo, however, this credit system has endured during the project's lifetime. The focus of this Chapter is to understand how the credit system and other project statistics, such as the number of models completed by a volunteer, have become embedded in the project (both in terms of league tables on the project websites and within the community of volunteers that has formed around the project), and the roles they play for various participants (both project team members and volunteers) in the project.

The systems of statistics operate in a uniform fashion across the project (i.e. all volunteers are subject to the same policies regarding the attribution of statistics, for instance how much computational processing work is represented by a single credit and what counts as completed for the point of view of counting the number of models completed by a volunteer). Furthermore, the policies that govern how they operate have been largely unchanged over the lifetime of the project.

The core of my argument is that although, on the one hand, the mechanisms by which the statistics systems operate can thus be understood as largely invariant at all points of the project both spatially and temporally[29], their social meanings are nevertheless highly sensitive to context, in the sense that they have had disparate meanings across the contexts of different individuals' lived experiences of the project[30].

This argument is split across the following two Chapters as follows: this Chapter focusses on the project team members, in particular their initial expectations of how they might achieve their project's major goals and

[29] In this sense, the credit and statistics systems might be understood as examples of *immutable mobiles* (Latour 1987). These are socio-technical networks which are invariant as they move through space and time. Examples can include systems of measurement and measuring devices, maps, scientific theories, experimental protocols, and so on.

[30] As I explained in Chapter 4, subsection 4.1.1, the social meaning of things (objects, concepts, diseases etc.) can vary a great deal between contexts because their identities are to be located in the networks they make with other things, and the relations available to them to form these networks can vary across contexts (Mol 2002; Mol & Law 1994).

how this changed as a result of their experiences, and the role of the credit and other statistics systems in this. The following focuses on the significance of the credit system in the experiences of some volunteers in the project.

In this Chapter, I argue that the project team members launched the project with the twin aims of both producing scientifically significant results and engaging a large number of members of the public. Furthermore, they commenced the project with a series of beliefs about volunteers and how to maintain their engagement and motivate them to contribute more to the project, namely that they would be engaged by greater education about the science underlying the project, and by the project's screensaver. These beliefs were based on a particular form of the *deficit model* (see Chapter 2, subsection 2.2.1).

I then argue that during the course of their involvement with the project and the volunteers, the team members find that these prior assumptions are challenged and confounded. In order to provide an account of how they have engaged and retained a large number of volunteers, they undergo a process of *social learning* (see Chapter 2, subsection 2.2.1), in which they seek to understand the processes by which the volunteers decide whether to remain involved with the project. This has resulted in a shift on the part of the project team members shift to conceptualizing the volunteers in terms of a different form of deficit model, in which volunteers are believed to be motivated primarily by the various statistics systems and league tables and thus exchanges between scientists and volunteers operate according to a market economy of knowledge. It will be shown how the team members have been influenced to adopt this belief for a variety of interacting reasons and, that furthermore, adopting this explanation for volunteer retention might be understood as a pragmatic move for the team members, in the sense that it is both a credible position (as it accords well with the reported experiences of others) and that it allows the team members discursively to construct a boundary between themselves and the volunteers, and thus maintain their authority in the project.

The main body of this Chapter starts by discussing the history of climate change modelling with the aim of contextualising (historically, politically, scientifically) the climateprediction.net project (Section 6.2). By considering how climate modelling has developed scientifically (and highlighting technical challenges to producing accurate forecasts of future climate change) and situating it in the context of growing concern about the impact of climate change (including how climate modelling has contributed to this) and subsequent contestations of policies to mitigate future climate change (including the ways in which technical limitations of climate models can be used to argue against the implementation of such policies), the main motivations for the founding and funding of climateprediction.net can be understood. Section 6.3 then details the running of the project, including a description of its scientific work and features for volunteers.

By considering the contexts discussed in Section 6.2, and thus by locating the interests of these tea members in these broader socio-technical networks, I argue that the credit system becomes very important to team members as they attempt to pursue two interests in particular, namely to run a distributed project that 1) produces scientific results of interest to the cutting edge of climate science; and 2) recruits and retains a large number of volunteers. In Section 6.4 (particularly subsections 6.4.1 and 6.4.2), by drawing on publications written by project team members, supplemented by interviews with them, I will argue that the credit system is important for them as it forms the basis of their understandings of the volunteers in terms of how they contribute to the team members' pursuit of both 1) and 2) but that also allows them discursively to reinforce the boundary between them and the project volunteers to maintain their scientific authority – in these accounts, the team members assert that project volunteers in general are motivated to participate primarily by the accumulation of credits and the social status this brings in the volunteer community.

However, I will also argue that the value of the credit system emerged during the course of the project: prior to the project's launch, the team members regarded it as being of trivial importance, instead suggesting that other features of the project would enable the project to achieve goals 1) and 2), and I will attempt to chart the circumstances under which the team members increasingly came to view the credit system as very valuable in the pursuit of their interests as the project unfolded and their understandings of the volunteers changed. Subsection 6.4.3 offers some interesting comparisons with Galaxy Zoo regarding how the respective project teams' interests emerged (both before and during the projects), how their expectations and understandings of the volunteers changed over time, and how this impacted on their decision-making regarding acknowledging volunteer contributions: these will be drawn out in the Conclusion.

6.2 climateprediction.net in context

In this Section, I will consider the broader contexts into which the climateprediction.net project emerged, in particular that, by the end of the 1990s (when the project was founded), two key challenges faced the field of climate change modelling. One was to improve scientific understandings of likely future changes in climate change modelling in order to strengthen policy-making processes, and the second was to build and secure public support for the field, via an acceptance that:

1) Climate change is a likely consequence of human activity and a threat to human well-being;
2) Governmental and intergovernmental action as a result of policymaking is required to mitigate this threat; and

3) Improved understandings of likely future climate change, through further work on climate change model, are critical for this policymaking.

In subsection 6.2.1, I will sketch out a history of post-World War II climate change modelling, explaining how improved (and improving) computational technologies afforded great improvements in computer-based weather forecasting to the point of vastly superseding previous methods, and how longer-term forecasts of climate change could be developed from weather models. This Section then considers how increased concern about long-term man-made climate change in the 1960s and 1970s led to greater institutional and political support for climate change modelling (and, in turn, how the output of climate change models contributed to growing awareness of the issues), and how the output of such models subsequently had an increasing influence over the policymaking process.

However, the field of climate change modelling has not simply just carried all before it: in subsection 6.2.2, I will explain some of the variety of technical and scientific limitations to producing accurate climate forecasts that have emerged over the past decades. At the same time, a number of vested interests have felt threatened by policies to mitigate climate change, leading to an increasingly powerful lobby to oppose such policies, which first began to break through into public discourse during the 1980s. This lobby has exploited the uncertainties arising in climate change modelling to attempt to undermine the scientific basis for policy making. Thus the two key challenges described above can be understood.

6.2.1 Weather forecasting and the development of computational models for climate change[31]

Computational approaches to modelling climate change grew out of the development of computational methods for weather forecasting. Until World War II, weather forecasting was done by hand by teams of forecasters who would use data about recent weather conditions and their expert knowledge to predict short-term changes in the weather. The science of meteorology received public support and funding as a result of its usefulness for a range of purposes, including military, agricultural, and seafaring. Due to the vast number of factors that were known to impact on weather, the contemporaneous methods only achieved limited success in accurate forecasting. However, early attempts to develop numerical methods, using differential equations modelling atmospheric circulation, met with failure as equations proved analytically intractable. As a result of this, it became clear that weather forecasting could not be performed by solving systems of equations, but instead should be approached using iterative modelling to find numerical approximations to these equations, where a set of initial conditions describing the

[31] This subsection broadly follows the history of climate change modelling set out in Weart (2010).

current state of the weather would be fed into a system of equations that would describe the change in weather over a fixed period of time (call this t). If x is the current point in time, then the output of these equations would be taken to represent the state of the weather at time $x + t$. This output would then be fed back into the equations, and the subsequent output would be taken to represent the state of the weather at time $x + 2t$, and so on. In order better to approximate the differential equations, t needs to be as small as possible (the differential equations are the limit of the iterative system as t tends to zero): as t decreases, the number of iterations required to produce a simulation for a fixed period of time increases proportionately (for instance, if t is halved, then twice as many iterations are required). Furthermore, in order better to simulate the weather, as many relevant parameters as possible need to be included in the system of equations. These two requirements for producing a realistic simulation of the weather meant that producing such a simulation by hand was an extremely time-consuming process, and thus many scientists came to the conclusion that this approach would not prove feasible.

However, in the early 1950s, scientists started to explore the possibilities for developing simulations of atmospheric circulation in light of advances in computational technologies that had come about during World War II. By 1955, scientists had developed models for regional weather forecasting sufficiently accurate and fast that the US Weather Bureau set up a unit that issued computer-based weather forecasts. At around the same time, the first model that simulated global – as opposed to regional – weather was developed. As a result of these successes, the US mathematician, John von Neumann, secured US government funding to launch a programme in 1955 to develop *General Circulation Models* (GCMs), simulating the entire global atmosphere.

Over time, modellers incorporated more parameters into their models. This was due both to increased computational power, and increasingly reliable and comprehensive real-world observational data of these parameters. Increased computational power also allowed for more fine-grained models, both in terms of reducing the length of time steps (t, as above) and at a higher level of resolution.

In the mid-1960s, some modellers turned their attention to climate rather than weather, using their GCMs as a basis to develop climate models for a seasonal or annual basis. There was a growing perception that GCMs could be used as a basis for experiments, in the sense of varying parameters or structures of a numerical model to see what might happen to the Earth's climate. These developments coincided with reports of observations of increases of carbon dioxide in the atmosphere, and of other greenhouse gases in general. As a result, some modellers began to ask what might happen to the Earth's climate in the future if the level of carbon dioxide continued to increase. In 1967, results were published stating that the GCM at the US Weather Bureau suggested that a doubling of carbon dioxide would produce an increase in global mean temperature of approximately 2^{O} C.

By the early 1970s, following a series of extreme weather events and the OPEC energy crisis, global warming became an item of serious concern for many members of the public in both the USA and Europe. This gave impetus to the further development of GCMs for climate change modelling, with the aim of producing more reliable forecasts for government policy making. In 1979, an investigation mandated by the US President's Science Adviser and conducted under the auspices of the US National Academy of Sciences, resulted in a report asserting with a high level of confidence that the doubling of carbon dioxide in the atmosphere would result in global warming of three degrees, give or take 50%.

The models to-date treated the oceans as fixed, but there was a growing awareness that circulation in the oceans also impacted upon climate in the sense that the oceans could absorb large quantities of atmospheric gases, move them around, and release them elsewhere. In order to improve climate simulations, some scientists set to work during the 1970s to develop models of ocean circulation and couple these models with models of atmospheric circulation. The results of these models were largely in line with the forecasts of atmospheric-only models regarding what might happen as a result of the doubling of carbon dioxide. To this day, coupled ocean-atmospheric circulation models remain the gold standard of climate change modelling.

By the late 1980s, as a result of the consistency of climate change forecasts and widespread public concern about global warming encouraged by reports in the popular media (Weingart *et al*. 2000), climate change was being taken seriously by governments and was the focus of a large global network of researchers based in many prestigious institutions (both Universities and other organisations like NASA). For instance, in 1996, the International Panel on Climate Change produced the first of its five-yearly reports on the state of the art in climate change research and policy proposals. This was followed in 1998 by the Kyoto conference, where governments agreed targets for reductions in emissions of greenhouse gases over the coming decades.

6.2.2 Challenges to climate change modelling: Technical and social

However, the broad consensus amongst the results from the various coupled GCMs masked a great deal of differences between models. As I discussed above, the models developed in complexity according both to what could be feasibly simulated on the available computational power and to the real-world observational data. Different teams of climate model-builders dealt with these constraints in disparate ways, highlighting a number of different sources of uncertainty in building climate models, in the sense that at various points in the model-building, it may not be obvious which of a range of options is the best one to take. Indeed, as Lenhard & Winsberg (2010) argue, the analytic intractability of the equations involved, and the fact that they are providing forecasts for the coming decades (and thus, the accuracy of their forecasts cannot be verified for some time to come), mean that it would be impossible to single out one particular climate model as being definitively better than the others. Instead, they argue, that the plurality of models is here to stay.

Parker (2010) has identified three main types of uncertainty, which have been a persistent feature of climate change modelling. These are:

1) **Structural uncertainty,** which refers to uncertainty about the form of the modelling equations, i.e. of which variables should a particular quantity be represented as a function? For instance, should a particular quantity be represented as a function of variable *x* only, or of variables *x* and *y*, or some other combination of available variables? This uncertainty can arise because some processes are still not well-understood theoretically, or because some equations might prove too computationally demanding to calculate and need to be simplified;
2) **Parametric uncertainty,** which refers to uncertainty about the values that should be assigned to different parameters within a system of equations;
3) **Initial condition uncertainty,** which refers to uncertainty about the state of the atmosphere that should be used as the starting point for the model. There are many challenges to accurately measuring the temperatures, pressures, wind speeds etc. at various points in the atmosphere. For instance, in measuring temperature at a particular point in the atmosphere, should a single measurement be used or should an average be taken of multiple observations over a period of time, and if so, what period of time?

Differences in the ways in which these types of uncertainties are addressed can potentially result in very different climate forecasts. For instance, in the case of structural uncertainty, if one model expresses a quantity as being a function of one particular variable and another model expresses it as a function of a different variable, then the value of that quantity might change in very different ways as each model progresses. And in the cases of parametric and initial condition uncertainty, it has been found that climate models can be sensitive to very small perturbations of parameter values and initial conditions respectively.

There are at least two major implications of this. The first is that, due to a model's potential sensitivity to even very small variations, the results of a single run of a climate model with a fixed set of parameters and initial conditions has only very limited value or significance. One approach is to run an ensemble of climate models, in which multiple runs are performed using a variety of models. Ensemble methods had been introduced into weather forecasting in the early 1990s, and by the mid-1990s, work was well underway to develop ensembles for climate change modelling as well. In the cases of some ensembles, these models vary structurally and are known as *multi-model ensembles*[32]. Other types of ensembles are known as *perturbed physics ensembles*, and involve running multiple models whose parameter values or initial conditions vary over a plausible space of values (i.e. these values are perturbed in the sense that small variations are introduced). climateprediction.net is an example of a perturbed physics ensemble.

The second major challenge that has been fuelled by this uncertainty is that it has given climate change sceptics a handle for challenging the scientific consensus around climate change. Fox Keller (2011) has suggested a range of motivations for sceptics, including a resistance to government regulation. For instance, many of the policy proposals made to reduce carbon dioxide emissions involve the regulation of industry, or costs to individual citizens and businesses (for instance, taxes on petrol), and thus many powerful groups are disposed to viewing climate science as a threat to their interests. That representational uncertainty existed in climate change modelling allowed sceptics to argue that climate models did not, in fact, represent a solid basis for concluding that climate change was happening or that it was related to human activity[33], and this was increasingly reported in sections of the mainstream press, thus undermining public support for climate science and policies to combat climate change. By 1996, a discernible shift had occurred in statements by governments and international organisations about the prospect of future climate change away from the confident pronouncements of the late 1980s to a more cautious tone (Weart 2010). As I discussed in Chapter 2, subsection 2.2.1, the late 1990s saw an increasing shift in the political culture in Europe and the US towards expectations that scientists would seek to engage the public in their work: this, combined with the growing concern about the potential immediacy and severity of the threat of climate change, and the growing movement of sceptics countering such concerns, meant that climate science in particular was facing a critical challenge to engage members of the public and gain their support.

Thus, by the late 1990s, climate science needed both to reach out to the public, and to devise ways of running ensembles of models. It is into this context that climateprediction.net emerged.

6.3 The climateprediction.net project

climateprediction.net was founded in 1999, funded by the *Natural Environment Research Council* (NERC). Its primary goal is to run ensembles of climate models on a distributed network of computers owned by members of

[32] One example of this is the *Coupled Model Intercomparison Project* (CMIp3) (Solomon *et al*. 2007).
[33] Climate change sceptics have picked up on other features of climate modelling as well in order to try and support their case. For instance, it was found that climate models drifted slowly over time into unrealistic patterns and ad hoc adjustments were made to correct this (Shackley *et al*. 1999), leading to accusations that climate scientists were fiddling their models to arrive at the politically correct answer that climate change is happening.

the public to quantify the impact of parametric and observational uncertainties on forecasts of climate change. It was launched to the public in August 2003.

This Section considers the circumstances under which this project was founded, and how it has since operated. It might be assumed that the project emerged unproblematically as a solution to the challenges described in Section 6.2. However, there were a range of alternative approaches to addressing these. For instance, rather than ensemble approaches, funding and support could be given to research into improved observations of the various physical phenomena so that parameterizations and initial conditions of climate models are more accurate or to improving computational technologies to support more sophisticated models. Even if ensemble approaches were to be supported, why not invest more in supercomputers to run them rather than getting the public involved?

Thus, climateprediction.net should be understood not simply as responding to the prevailing contexts of the time, but that key individuals involved actively helped to create the socio-technical space into which it emerged. This is considered in subsection 6.3.1 and, furthermore, through their actions it can be understood how members of the initial climateprediction.net team, and in particular Myles Allen, became personally committed to demonstrating that the approach of running ensembles of climate models on a distributed network of computers owned by members of the public could make a significant scientific contribution to the field of climate modelling.

At around the same time, the NERC launched its *Coupled Ocean Atmosphere Processes and European Climate* (COAPEC) programme to improve understandings of climate change forecasting, issuing a call for projects to apply for funding. Fully aware of the need to build public support for the field, NERC saw climateprediction.net as a way of engaging the public in climate science and provided funding, in the hope of it making a scientific contribution and in the expectation of recruiting a large number of volunteers.

Thus, subsections 6.3.1 and 6.3.2 together can be seen to provide a foundation for understanding that the two main aims of climateprediction.net at its launch were:

1) To produce genuine scientific research, to realise the vision of the projects' founders for the scientific potential of its approach to ensemble methods; and
2) To recruit a large number of members of the public to actively participate in the project[34].

Section 6.3 concludes with a description of the work of climateprediction.net, and its features and policies for members of the public. This provides essential background knowledge for the discussions in the subsequent Section (6.4).

[34] At this point, I would like to stress that these goals are not necessarily related – a significant scientific output does not necessarily require a large volunteer base, and an increased user base does not necessarily lead to an improved scientific output. For instance, the vast majority of the project's scientific output has been generated by a small minority of volunteers.

6.3.1 *'Do-it-yourself climate prediction'*[35]

It became apparent that running ensembles of coupled GCMs would require vast computational power, of a scale that was only available at the time in a handful of supercomputing centres around the world. Inspired by the example of early public distributed computing projects such as SETI@Home, Allen (1999) suggested that the ensemble approach to GCMs might be carried out using a similar approach. To set about realising this vision, Allen put together a team of scientists from the Atmospheric, Oceanic and Planetary Physics Group at the Department of Physics, University of Oxford and the Hadley Centre at the Met Office.

This team formed the core team of the climateprediction.net project, with Dr Allen as the project's Principal Investigator. As the individual who had proposed this approach, Dr Allen became personally committed to demonstrating its potential for success, and this has been enshrined as one of climateprediction.net's key objectives from the very beginning. For instance, in a paper which preceded the public launch of the project, it was written that one of the major *'overall objectives'* of the project was to *'demonstrate the potential of distributed computing for Monte Carlo ensemble simulations of chaotic systems'* (Stainforth *et al.* 2002, p. 32).

Even after the project launch, the importance of demonstrating the success of this approach remained a critical motivation for Allen (and other members of the project team). This is illustrated by the eventual awarding of a patent in 2006, through which the team can be understood as establishing their priority for the conception and development of such a method. The patent was awarded for *'A method of generating short-, medium-range and seasonal-timescale weather or climate forecasts by running an ensemble of computer models on a distributed computing system and network'* (Allen *et al*. 2006, p. 1) and that, furthermore*, 'a large number of models can be distributed to such personal computers (for instance owned by the general public)'* (*Ibid.* p. 8).

Thus, it can be understood that the success of the project in terms of its scientific output was seen by the core project team as a critical goal for the project right from the outset. Nevertheless, in order for Myles Allen's vision to be realised, the team needed to secure funding in order to be able to develop the project for it to be launched to the public.

6.3.2 Institutional support for climateprediction.net

The main source of funding for climateprediction.net came from the Natural Environment Research Council (NERC), both as part of the NERC's Coupled Ocean Atmosphere Processes and European Climate (COAPEC)

[35] Allen (1999, p. 642).

programme and of its e-Science programme[36]. COAPEC ran from 1999-2006. Its overriding objective was to advance understandings of the impact of the coupling of climate models involving the atmosphere and the Atlantic Ocean on climate forecasting running on seasonal to decadal timescales. climateprediction.net was one of sixteen projects successful in the First Round of funding. The project was awarded funding for two reasons in particular. The first was its potential for building public support for research into climate change. Given that this round of funding was in the context of a political culture that called for greater public engagement with science, and in particular the critical importance of public support for climate change research for policy purposes, it is no surprise that the NERC was particularly keen to award some of the COAPEC funding to projects with a strong public engagement element.

As a result, climateprediction.net was awarded funding as the public outreach arm of the COAPEC programme, as the following extract from the project's final report explains:

> *'One of the key aims of climateprediction.net is to dispel the mystery that surrounds climate modelling and prediction through direct public involvement in a front-line climate research project. The public awareness of this project has provided a very public* "shop window" *to the kind of research carried on within the COAPEC programme'* (NERC 2007, p. 4)

The second major reason for awarding funding to climateprediction.net was to explore the possibilities afforded by public distributed computing for running ensembles of coupled climate models to forecast climate change over the coming decades:

> *'[The project] aimed to test the concept of running full-climate-resolution atmosphere-ocean general circulation models using idle time on personal computers'* (NERC 2007, p. 4)

Similar reasons led to the awarding of funding to climateprediction.net under the NERC's e-Science programme. This programme was part of a broader national UK e-Science programme. Through the national programme, approximately £250 million was provided to support the development of computing infrastructure to support distributed, interdisciplinary, data-intensive scientific research, of which £14.2 million was allocated to the NERC. The NERC funded eight projects involving collaborations between computer scientists and environmental scientists. These included climateprediction.net, which was allocated funding from 2002 to 2010. As with COAPEC, climateprediction.net was viewed as the public outreach arm of the e-Science programme.

It can therefore be understood that the climateprediction.net team were driven to build and run a project that both:

[36] In addition to the NERC funding, climateprediction.net has received a number of smaller grants from other organisations and institutions for various reasons. For instance, Microsoft provided funding to help improve the middleware and to facilitate an interface through which interested scientists might be able to access the data generated by the project, whilst the Nuffield Foundation gave a grant to help the project develop educational materials for use in schools.

1) **Resulted in a scientific output that made a significant contribution to the advancement of quantifying uncertainty in forecasts of climate change.** In this respect they were motivated both by the conditions under which they were funded by the NERC, and by the personal commitment of Dr Myles Allen towards realising his vision of ensemble climate change models being run using public distributed computing; and
2) **Recruited and retained a broad base of volunteers,** in order to fulfil the NERC's objectives for the project, and for the scientists themselves to build public support for the discipline in which they work.

Furthermore, it can be seen that these motivations existed prior to the public launch of the project, and persisted during the project as the project would be assessed at the conclusion of both COAPEC and the e-Science programme against its initial aims, and as the team sought to develop and demonstrate the distributed approach to ensemble GCMs.

6.3.3 The running of climateprediction.net

The project originally intended to run between one and two million versions of a climate model, which simulated future climate, each using different estimates for the model's parameters, to assess the level of variation generated by the model in predictions of future climate change (Stainforth *et al.* 2002). The project began in 1999 following Myles Allen's article in *Nature* discussed above in subsection 6.3.1. After alpha and beta testing phases in late 2002 and spring 2003, climateprediction.net was launched to the public in September 2003. Initially, the project did not use the BOINC middleware, but instead used middleware developed by the project team. However, in August 2004, the project moved to using BOINC, and continues to use this middleware. Over the years, the number of active users stabilised at around 25 000-40 000, and as of November 2011, there are approximately 26 000 active users.

The core science team were and are based at the University of Oxford in the Department of Physics, supplemented by a support team based mainly at the following institutions:

- *Oxford e-Research Centre*, University of Oxford. Team members based here are involved with the technical side of writing and developing the project software, and maintaining the project data servers; and
- *The Hadley Centre* at the UK Met Office. Team members here are responsible for the development of the climate models used in the climateprediction.net experiments.

The main scientific focus of climateprediction.net At the time I conducted my case study (2008-2009), there had been two phases of the project. The first, preliminary, phase was launched in 2003 at the time of the general public launch in order to study sensitivity to various perturbations in parameters, and the aim was to contribute to the design of the main climateprediction.net project, which was to be eventually launched in 2006.

The first phase involved the *HadSM3* model, developed by the UK Met Office, which simulates 45 years of climate. This is a comparatively simple General Circulation Model (GCM), in that it treats the ocean as fixed, rather than dynamic (i.e. it is not a coupled atmospheric-ocean model). Such a model is known as a *slab* model. As explained above, climate models can be highly sensitive to perturbations parameter values, even to the point that a very small perturbation can result in a model no longer being realistic and instead showing chaotic and unrealistic results. The aim of this first phase was to identify suitable ranges of parameter values to form the basis of the main, second phase of the project.

The second phase of the project was to provide realistic forecasts of future climate change over the period 2000-2080 by conducting millions of runs of a climate model that covered both the space of realistic parameter values identified by the first phase, and also covered a space of realistic initial conditions (i.e. dealing with both the parametric uncertainty and initial condition uncertainty discussed in subsection 6.2.2). The model used in this phase was the *HadCM3* model, also developed by the UK Met Office, which is a coupled atmospheric-ocean GCM. This model covers a simulation period of 160 years: it runs from 1920-2000 in order to provide a hindcast in order to see whether a particular combination of parameters accurately models the known climate over this period and, if it does, then the model runs from 2000-2080 to provide a forecast of future climate change. One climateprediction.net model is known as a *work unit*. A typical work unit for climateprediction.net is substantially longer than for almost any other BOINC-based project. For instance, a work unit for SETI@Home or Einstein@home *'can typically be completed in a matter of hours'* (Christensen *et al*. 2005, p. 9), whilst, at the time of the launch of the BBC project, it was expected that the 160-year HadCM3 model would take approximately four months to run on an average volunteer's computer.

This second phase was in conjunction with the BBC as part of its *Climate Change Season*. The BBC worked together with climateprediction.net on the *BBC Climate Change Experiment*, which was launched in February 2006 with a television programme explaining the goals of the project and encouraging viewers to enrol. Approximately 300 000 volunteers signed up through a web page on the BBC's own website, and the results were presented in a follow-up programme in 2007. After this, BBC ceased its involvement in climateprediction.net, but the project continues to recruit volunteers to run this 160-year forecast, in order to provide more refined forecasts[37], and further results have been presented in academic journals (for instance, Piani *et al*. 2005; Stainforth *et al.* 2005). By 2007, climateprediction.net had become well-established as a credible approach to tackling ensemble approaches to climate forecasting (Collins & Knight 2007).

Smaller subprojects In addition to its main research project, a number of smaller projects have been conducted through climateprediction.net, including:

- ***Geoengineering experiment.*** This has been running since October 2008, and seeks to forecast the impact of a range of alternative courses of action which might mitigate climate change in the future by running simulations of climate change under each course of action from 2000-2080;
- ***Millennium experiment.*** Using a speeded-up version of a well-established climate model, this seeks to simulate climate change from approximately 800AD until around 1900AD to assess the impact of various factors (volcanic and solar activities, land-use changes, and changes in ocean circulation patterns) on climate change over this period. The project was launched in 2010;
- ***weatherathome.*** This seeks to understand the impact of climate change on weather patterns by simulating the weather in various regions of the world under different climate change scenarios. This is a new project. Initially three regions of the globe are being targeted (USA, Europe, and southern Africa), with the intention of rolling the project out across the whole globe.

The volunteer experience of climateprediction.net climateprediction.net uses BOINC (see Chapter 1, subsection 1.1.1) and thus there are a number of features for project volunteers. The first is an interface through which the volunteer can communicate with the project and choose whether to participate in the main climateprediction.net experiment (and if so, whether to run the 160-year coupled GCM or the 45-year slab model) or to participate in one of the smaller other experiments. There is also a screensaver which displays to a volunteer a visualisation of the progress of their climate model (see Figure 6.1).

On the project website itself, there are a number of pages (see Figure 6.2 for an image of the website's homepage). One is the *News* section, where project team members can give updates of various issues relating to

[37] It should also be noted that the first phase of the climateprediction.net project is still ongoing, as the space of potential parameter values has not yet been fully-explored. The BOINC interface allows volunteers to choose whether to run this, simpler 45-year model, or the full 160-year coupled GCM.

the project (e.g. the launch of new sub-projects, publication of results from the project, technical issues currently being experienced by the project). There are also pages which seek to explain the science that underlies the project (such as how climate models operate, how the ensemble approach to climate modelling works, and how the project itself has been designed).

A further, important, feature is the project forums (see Figure 6.3). If climateprediction.net volunteers have difficulty with installing or running BOINC, or encounter difficulties with their climateprediction.net model, they can seek advice on these forums. These forums also provide the main arena for project participants to interact, and for those involved with the running and administration of climateprediction.net to communicate with the volunteers. Some are labelled to encourage participants to share problems they have, and offer opportunities for participants to offer encouragement or technical advice to others; others have been set up with the intention of participants sharing and discussing results from their models, or to discuss general issues relating to climate science. The responsibility for maintaining order lies with a group of moderators, appointed from amongst regular forum contributors.

Another feature of climateprediction.net is that it, in common with other BOINC-based projects, records statistics about the quantity of model data processed by project volunteers. Three key indicators are:

- **Total credit,** which measures the total quantity of data processed on an individual's computer for the entire duration of their involvement in the project. In climateprediction.net, credits are awarded for each model-year completed on a volunteer's computer;
- ***Recent Average Credit*** **(RAC),** which is a weighted average of the credit awarded to a volunteer, with the weighting given to the credit decaying exponentially over time, so that the greatest weighting is given to the credit most recently awarded; and
- **The total number of 160-year HadCM3 models completed.** It should be noted here that the decision was made that a model should be considered completed when it has reached 2050 in the simulation, rather than 2080 which is the simulation's end-point.

These are very often a source of pride for project volunteers, who will sometimes report their statistics in a signature placed at the end of forum posts. Furthermore, volunteers can group together and form teams, with one volunteer being designated as the team leader, and very frequently, these teams have websites and forums of their own. On the climateprediction.net website, there are lists ranking both individual participants and teams according to credits acquired, or RAC (see Figure 6.4).

It should be noted at this stage that, although the project only started using BOINC in August 2004, a number of the features listed above have been present since the public launch of the project, in particular the online forums and the credit system (including the team element and league tables).

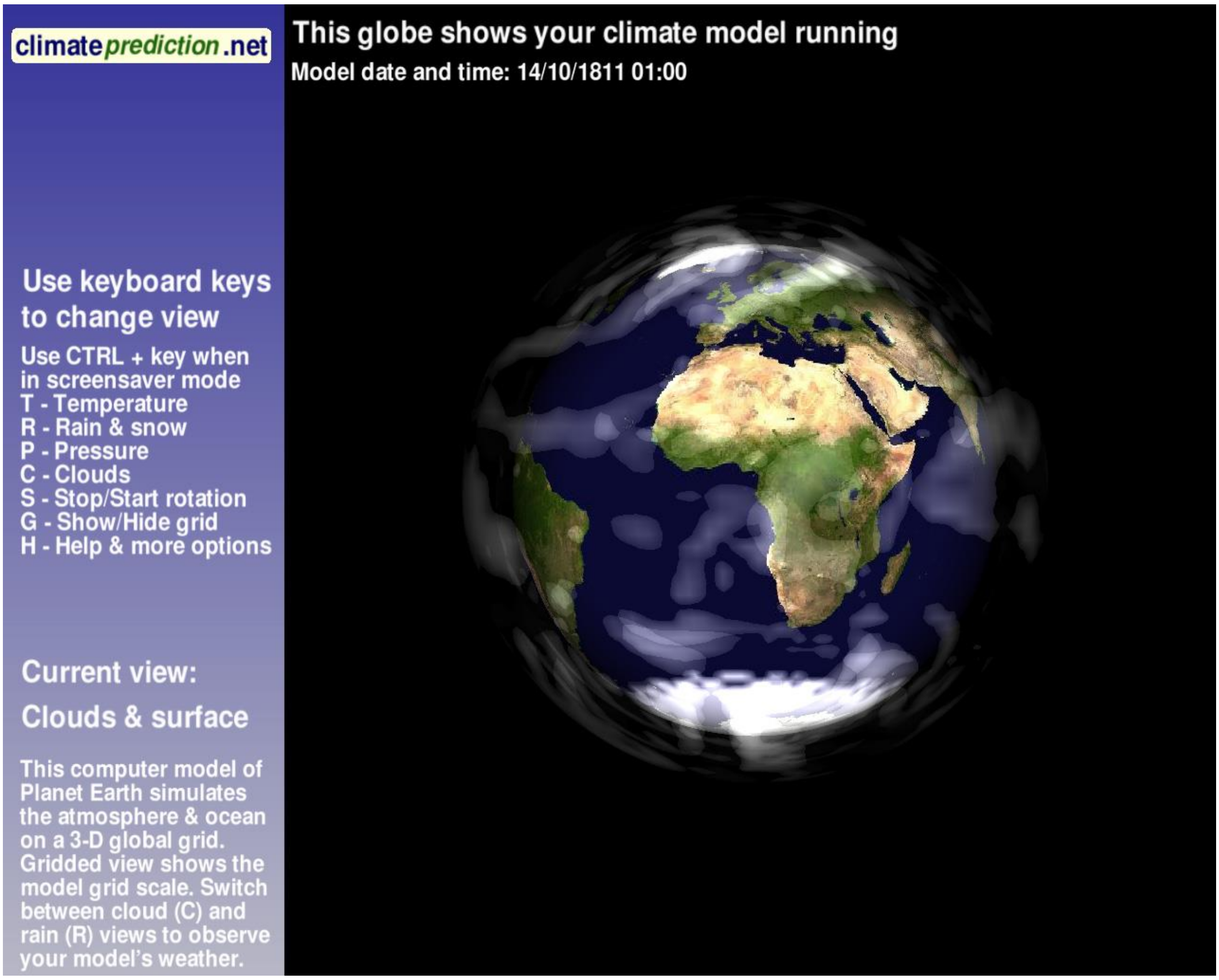


Figure 6.1: The climateprediction.net screensaver, *showing the status of a volunteer's model (image taken from http://www2.sunysuffolk.edu/mandias/global_warming/climate_models_accuracy.html (accessed 27 November 2011)).*

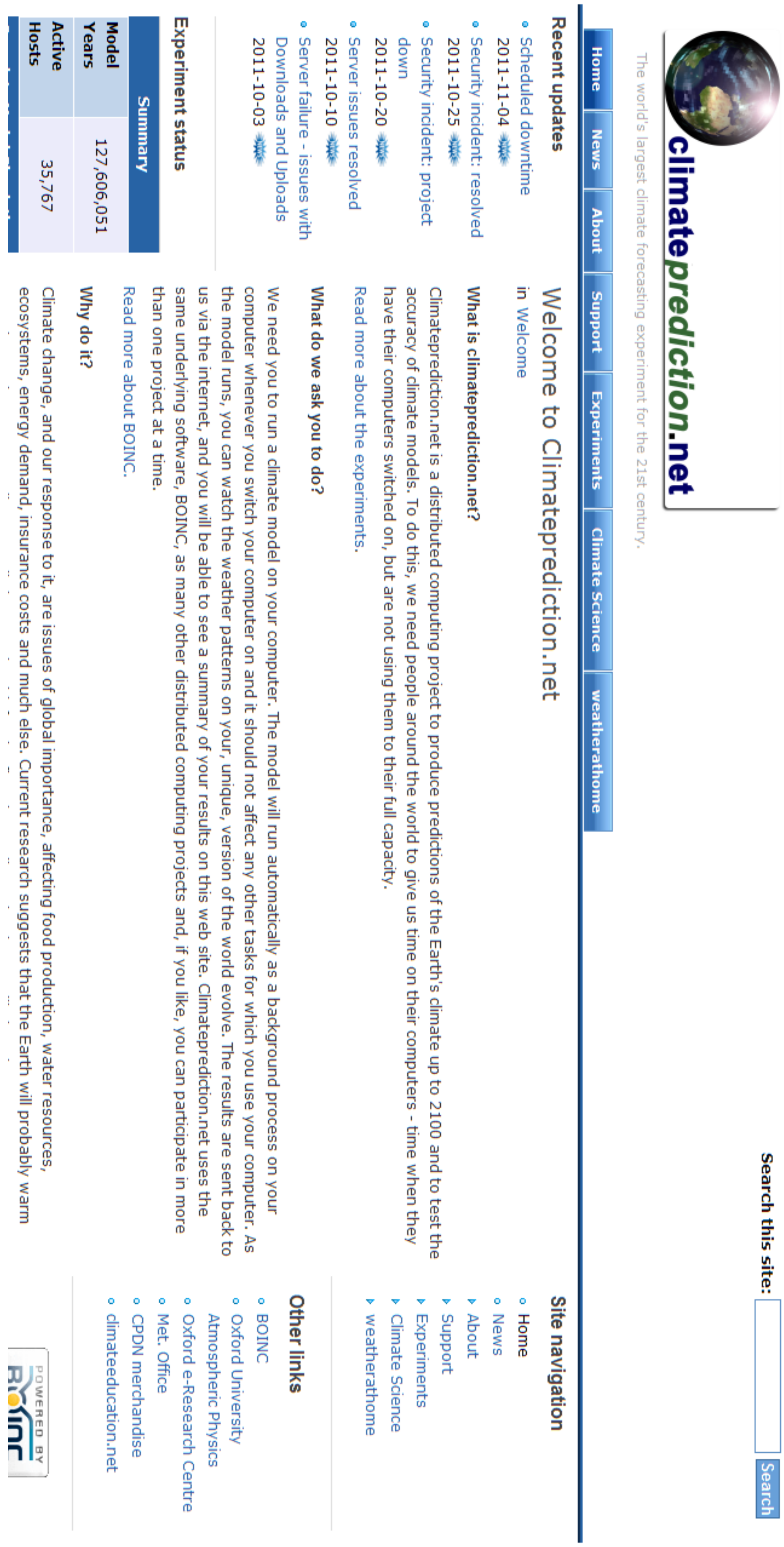
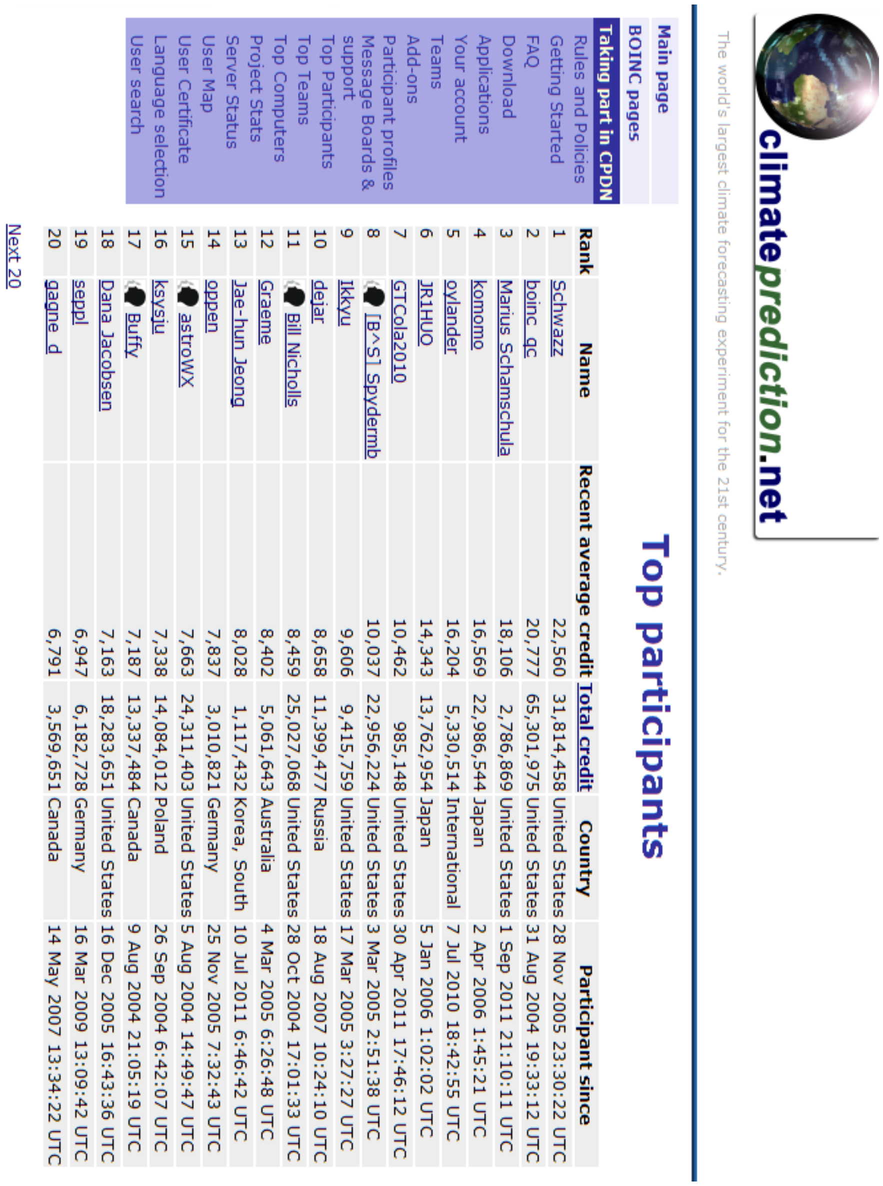


Figure 6.2: The climateprediction.net homepage *(http://climateprediction.net (accessed 28 November 2011)).*

Figure 6.3: The forums on climateprediction.net *(http://climateprediction.net/board/index.php (accessed 28 November 2011)).*



climateprediction.net

The world's largest climate forecasting experiment for the 21st century.

Main page
BOINC pages
Taking part in CPDN
Rules and Policies
Getting Started
FAQ
Download
Applications
Your account
Teams
Add-ons
Participant profiles
Message Boards & support
Top Participants
Top Teams
Top Computers
Project Stats
Server Status
User Map
User Certificate
Language selection
User search

Top participants

Rank	Name	Recent average credit	Total credit	Country	Participant since
1	Schwazz	22,560	31,814,458	United States	28 Nov 2005 23:30:22 UTC
2	boinc_qc	20,777	65,301,975	United States	31 Aug 2004 19:33:12 UTC
3	Marius Schamschula	18,106	2,786,869	United States	1 Sep 2011 21:10:11 UTC
4	komomo	16,569	22,986,544	Japan	2 Apr 2006 1:45:21 UTC
5	oylander	16,204	5,330,514	International	7 Jul 2010 18:42:55 UTC
6	JR1HUO	14,343	13,762,954	Japan	5 Jan 2006 1:02:02 UTC
7	GTCola2010	10,462	985,148	United States	30 Apr 2011 17:46:12 UTC
8	[B^S] Spydermb	10,037	22,956,224	United States	3 Mar 2005 2:51:38 UTC
9	Ikkyu	9,606	9,415,759	United States	17 Mar 2005 3:27:27 UTC
10	dejar	8,658	11,399,477	Russia	18 Aug 2007 10:24:10 UTC
11	Bill Nicholls	8,459	25,027,068	United States	28 Oct 2004 17:01:33 UTC
12	Graeme	8,402	5,061,643	Australia	4 Mar 2005 6:26:48 UTC
13	Jae-hun Jeong	8,028	1,117,432	Korea, South	10 Jul 2011 6:46:42 UTC
14	oppen	7,837	3,010,821	Germany	25 Nov 2005 7:32:43 UTC
15	astroWX	7,663	24,311,403	United States	5 Aug 2004 14:49:47 UTC
16	ksysju	7,338	14,084,012	Poland	26 Sep 2004 6:42:07 UTC
17	Buffy	7,187	13,337,484	Canada	9 Aug 2004 21:05:19 UTC
18	Dana Jacobsen	7,163	18,283,651	United States	16 Dec 2005 16:43:36 UTC
19	seppl	6,947	6,182,728	Germany	16 Mar 2009 13:09:42 UTC
20	gagne_d	6,791	3,569,651	Canada	14 May 2007 13:34:22 UTC

Next 20

Figure 6.4: The top 20 volunteers on climateprediction.net on 28 November 2011, as ranked by Recent Average Credit *(http://climateapps2.oerc.ox.ac.uk/cpdnboinc/top_users.php (accessed 28 November 2011)).*

6.4 The beliefs of climateprediction's team members regarding volunteer retention

It will be argued in this subsection that team members involved with running climateprediction.net who have had extensive experiences of interactions with volunteers generally attribute the success of the project in terms of volunteer retention primarily to the credit system (including the team system and the league tables), asserting that this has introduced a competitive element into the project which is a key motivating factor for volunteers. In particular, they do not distinguish between different volunteers (i.e. they do not suggest that whilst this may indeed be the main element that keeps some volunteers engaged in this project, other volunteers may instead continue their involvement for other reasons). As a result, these scientists and software engineers suggest that the credit system is the critical factor in their project's success in pursuing the goals of producing significant scientific results and engaging a very large number of volunteers.

However, I will argue here that this belief was not held prior to these scientists' and software engineers' involvement with the project (or, indeed, that this belief is shared by those who have had relatively little direct contact with the volunteers); instead, when the project was launched, they saw little reason for having a credit system as they anticipated that the educational possibilities and the screensaver graphics would be the key to volunteer retention.

I will suggest that what has happened here is that the scientists initially conceptualized the volunteers in terms a particular form of the deficit model (see Chapter 2, subsection 2.2.1), in which volunteers who have just joined the project are assumed to be interested primarily in helping science, but lack the necessary scientific knowledge and awareness of the processes of scientific reasoning that would allow them to understand the importance of climateprediction.net's work to climate science and how they are contributing to this through their participation in the project. Understood in terms of Mertonian norms, implicit in this conceptualization of the volunteers is the notion that they may be willing, in particular, to adhere to the norms of disinterestedness (as they are primarily motivated by the idea of advancing the pursuit of scientific knowledge) and communism (as they are willing to donate their own time and effort to science). By contrast, also implicit in this conceptualization, is that the volunteers may possess a deficit of the norms of universalism (as they may be ignorant of the relevant, already-established scientific knowledge) or organised scepticism (as the volunteers may not be aware of the *'empirical and logical criteria of judgment'* (Merton 1949, p. 276) that would allow them to evaluable scientific claims in a scientific manner).

I then suggest that, after they had let members of the public into the project and had interacted with many of them, the project team members' initial expectations were confounded. As a result, they reconsidered their assumptions about how the lay volunteers encounter and deal with science (see Chapter 2, subsection 2.2.1). This has resulted in a shift of their conceptualizations of the volunteers to another form of deficit model, one which implicates a view of the volunteers as having a deficit of the Mertonian norms of universalism and communism, instead invoking a market economy of knowledge underpinning exchanges between the volunteers and project team members (see Chapter 5).

It should be made clear at this point that in their discourse about project volunteers, the project team members do not distinguish between different volunteers; instead, in almost all instances (both before and after the project launch, and in both interviews and publications), the volunteers are usually portrayed as a homogenous group of people with respect to their motivations for participation.

6.4.1 How to retain volunteers? Part I: Screensavers and education

In papers published by members of the climateprediction.net team and in my interviews with team members, a number of different features of the project have been invoked at various times as playing an important role in the retention of volunteers, in particular the educational potential of the project in terms of volunteers learning about climate science, the screensaver that shows volunteers the progress of their current model, the statistics systems (including league tables and teams), and the online forums. In the argument that follows, I will chart the changes over time in terms of how the climateprediction.net team members understood the relative contribution of each of these factors to the task of volunteer retention.

Before the public launch of the project, the climateprediction.net team did not apparently anticipate that the credit system would be of any importance to the project, and it was not part of their strategy for volunteer retention. This can be seen in the following quotation from an interview with a member of the climateprediction.net team, who recalls that the decision to include a credit system at the public launch of the project was made at the last minute, and for no stronger reason than it being a feature of another major CCP:

> **Extract 6.1:** *'When we launched the original project pre-BOINC, we nearly didn't have a credit system. It was the night before the launch that we put together this credit system, that we decided since SETI has done this, let's follow them.'* (Interview cpdn-a/2)

Instead, the climateprediction.net team generally anticipated that the educational aspects of the project and the screensaver would be sufficient to maintain volunteer interest. For instance, in their reflections on their experiences in climateprediction.net, Christensen *et al*. (2005) stated that initially, they believed *'that the science and educational aspects, as well as the graphics, were sufficient to retain users' interest.'* (p. 4). Furthermore, it was felt that the credit system would be of little importance in this respect. The following extract from a paper published prior to the project's launch explains in more detail how they were hoping to retain volunteers with the screensaver and educational aspects of the project:

> *'[Keeping participants interested] requires that participants are actively involved in the experiment and feel a sense of ownership of their simulations. There is therefore a need for a dynamic, personalized web site and visualization software with which participants can view the progress of their simulation, as it develops…In the medium term there are also…possibilities for mini-research projects in primary, secondary and tertiary education'* (Stainforth *et al.* 2002, p. 35)

They further go on to explain that the visualisation should *'be both exciting and informative...It is anticipated that it will be a crucial part in involving secondary and tertiary educational establishments'* (Stainforth *et al.* 2002, p. 37). In other words, not only is the screensaver in and of itself important for volunteer retention, but it is also important in terms of helping to contribute to some of the educational aspects of the project.

Embedded in these statements are assumptions that suggest the project team originally conceptualised the volunteers, and the task of volunteer retention, in a form of the deficit model. Before going further with this analysis, I will first provide a brief recap of the deficit model and social learning.

The deficit model and social learning The deficit model is a model of the public understanding of science that has a number of key assumptions embedded in it. One is that scientific knowledge is universal, and has been arrived at by objective processes. Thus, under this view, dissent or challenges to this knowledge (or to a scientist presenting their knowledge claims) by lay members of the public (for instance, in the case of technoscientific decision-making, or advice about matters of public health) are seen as following from ignorance on the part of the dissenter about the relevant scientific knowledge, or ignorance about the correct processes of how this knowledge is generated, or as a result of irrationality on the part of the dissenter in the sense that they have allowed their emotions to cloud objective thinking, or are only concerned about their own personal self-interest (Wynne 2006). A corollary of this view is that if the lay public can be given the necessary education (taught the relevant scientific knowledge how to evaluate knowledge claims in a scientific manner) and persuaded to think objectively, then they would come to see the issue-at-hand as the scientist see it, and would thus no longer dissent but fully support the scientist. Here, the laypeople are seen as empty vessels to be filled unproblematically with scientific facts (Gregory & Miller 1998).

As discussed in Chapter 2 subsection 2.2.1, Wynne (1992) suggested that technoscientists who initially conceptualise members of the public in terms of the deficit model would find their expectations confounded, and would be forced to undergo *'social learning'* (p. 282), in which they would reflect on their assumptions and seek to come to an understanding of how laypeople evaluate scientific knowledge claims presented to them.

The deficit model and climateprediction.net team members The issue-at-hand for the project scientists is to ensure that volunteers continue participating in the project. As I explained above, volunteers can encounter technical difficulties in running the project middleware on their computers, which may deter their further participation in the project. Furthermore, they are able to stop participating in the project in favour of other CCPs. As a result, a volunteer may need persuading to continue with their participation in climateprediction.net, in the face of these incentives to stop.

The initial belief on the part of the project scientists that educational aspects of the project (both in terms of its use in educational institutions and in terms of the materials and information available on the project website), along with the dynamic screensaver, strongly suggests that these scientists initially conceptualised the task of motivating volunteers to continue participating in terms of a deficit model which sees lay people as lacking the necessary knowledge of scientific facts and scientific reasoning processes (both empirical and logical) to understand the project's scientific work. In this model, the education of laypeople about this knowledge is seen as the key to ensuring they choose to continue participating in the project, i.e. if only they possessed this knowledge then they would share the team members' understanding that the project makes an important contribution to the advance of scientific knowledge and appreciate why they should continue to participate. In this sense, then, it can be understood why the educational aspects of the project and the screensaver (in so far as it contributes to volunteer education) were perceived as important for volunteer retention. What has happened in the above extracts is that the dropping out by a volunteer from the project has been transformed into the failure on the part of the volunteer to take advantage of, or to fully appreciate, the educational materials made available to them.

This particular form of the deficit model is further invoked in the above quotations by the way in which it is seen as being linked to the notion that *'participants are actively involved in the experiment and feel a sense of ownership of their simulations'* (Stainforth *et al.* 2002, p. 35). The goal here is to encourage volunteers' interests to align with those of the scientist, in the sense of making the volunteer feel that they have a stake in the project and are responsible for a part of its scientific work.

Taken together, then, it can be seen that the educational aspects and the screensaver were initially regarded by climateprediction.net team members as the key to retaining volunteers, and embedded in this was the assumption on the part of the team members that the volunteers (and the task of their retention) could be conceptualised along the lines of a deficit model, which conceptualizes volunteers who enter the project as wishing to help science (Mertonian norms of universalism and communism) but lacking knowledge of scientific facts and reasoning processes relevant to their evaluations of whether to continue participating in the project. That this has been the case should come as no surprise: as Wynne and others have found, such a conceptualisation is common amongst scientists in general (e.g. Wynne & Felt 2006).

6.4.2 How to retain volunteers? Part II: Credits, teams, and social status

Soon after the project was opened to the public, however, these team members appeared to re-evaluate their perceptions, shifting to a different form of deficit model based on the view that the volunteers are primarily motivated to continue their participation in the project by the competitive acquisition of credits and the gaining of status in a community of volunteers by improving their league table rankings, rather than by helping the project to achieve its scientific goals. This belief is widely-reported by climateprediction.net team members[38]. For example:

> **Extract 6.2:** *'[Credits] became crucial, without it I don't think there would be any climateprediction project…if you don't have it, people won't do any work for the project at all'* (Interview cpdn-a/5)

It is also reflected in the following extract from a paper published in 2005, which suggests that many volunteers are motivated by the competitive element of credit accumulation:

> *'…they get* "credits" *for completing workunits. Many users enjoy competing with each other in terms of number of workunits finished and speed of computation of a workunit. Teams of computer users form and compete with one another.'* (Christensen *et al.* 2005, p. 4)

In particular, the project team members tie this motivation to the existence of the online forums, because it is through these forums that a community of volunteers has formed around the project, and it is within this community that the credit scores and league tables become meaningful, in the sense that the project scientists believe that they provide the foundation for this community's social structures and hierarchies. This is illustrated by the following extract from one of my interviews, where the online community is linked with the league table:

> **Extracts 6.3:** *'The other thing is probably the online community with the leaderboard, to keep people engaged, it's a very, very active community...making sure you have an active credit system to support the online community'* (Interview cpdn-a/2)

[38] It should be noted that there is one exception here. In a paper published by members of the climateprediction.net team in the journal *Advances in Geosciences* in 2006, both the screensaver graphic and educational possibilities of the project are credited as contributing to volunteer retention, along with the credit system, without privileging the credit system in this respect (Massey *et al.* 2006). This, therefore, seems more in line with the attitudes of the team members prior to the project's launch. This apparent anomaly, however, might be accounted for by considering the audience for this paper, and the purposes of the author. It was published in a journal whose primary readership are geoscientists, rather than the computer scientists and those involved with e-science, and who therefore are highly unlikely to have had experience of setting-up and running a CCP. Secondly, the main focus of this paper was not about experiences of interactions with the volunteers – the quote above is from the Introduction of the paper and this is the only place where the issue of volunteer retention is mentioned. Thus, the apparent anomaly could be explained not as dissent on the part of experienced climateprediction.net team members regarding the critical importance of the credit system for volunteer retention, but instead as part of a strategy for gaining credibility amongst an audience with no experience of CCPs by writing about strategies for volunteer retention in a way that accords with the expectations of this audience in order to persuade them that the authors' core arguments should be taken seriously.

This point was made more explicitly in an interview with another climateprediction.net team members, who suggested that those with a very large credit score were particularly respected by other volunteers on the forums, and were regarded by others as being especially authoritative on matters relating to the project and climate science in general[39].

Social learning on the part of the project team members So what brought about this change in the team members' beliefs about why volunteers are motivated to participate? In accordance with what other literature on the deficit model (e.g. Wynne 1992) might lead one to expect, it can be seen that initial expectations about the climateprediction.net project would be able to retain volunteers were indeed confounded as the team members engaged with the volunteers in various ways. At this point, I would like to stress that I am focussing on the beliefs of the team members, and thus am seeking to provide an account of their social learning in terms of the events they observed and their interactions with the volunteers, how they interpreted these, and external influences on their interpretations.

In terms of the contribution of education to volunteer retention, many of the team members were apparently surprised by how little interest the volunteers appeared to have in learning more about climate science. In the early days of the project, the following team member posted on the forum on a regular basis, and they describe their experiences in the following two quotations:

> **Extract 6.4:** *'When they saw me on the forum, I expected them to want to suck out as much science from me as possible, but they didn't seem interested'* (Interview cpdn-a/3)

And they went on to say:

> **Extract 6.5:** *'Going to the forum...it seems to be old-school socialising, people ask each other how they are, and make funny comments, so these people don't do this for some sort of pure science, they do it for some sort of fun and they like the community'* (Interview cpdn-a/3)

This is echoed by another interviewee, who stated:

> **Extract 6.6:** *'The volunteers are not asking us for scientific knowledge'* (Interview cpdn-a/4).

In other words, the team members anticipated that the volunteers would see this person's presence on the forum as an opportunity to learn more about the science, one that they expected the volunteers would be keen to exploit. However, they felt that this expectation was confounded, interpreting the volunteers' behaviour in terms of a lack of interest in climate science. Instead, they came to believe that the forum's success was due to volunteers seeking to associate with others because they enjoyed socialising online.

[39] The interviewee in question asked that I not record the interview – and thus no transcript of it exists – but did allow me to take notes paraphrasing what they said.

The other initial expectation on the part of the team members regarding volunteer retention was that the screensaver would play an important role. However, this expectation was also confounded very soon after the project launch. Initially, the project website included a contact form (which was abolished soon after the launch) where volunteers needing could send messages to the climateprediction.net team to ask for support for technical issues they were encountering. The topics raised by the volunteers suggested strongly to the project team members that the volunteers had little interest in the screensaver, relative to other features of the project, as the following quotation from an interview with a climateprediction.net software engineer explains:

> **Extract 6.7:** *'When I started, we had this support email, we had a contact form where people could make comments and ask questions, and it used to go on to our email account and I had about 500 emails a day from volunteers asking how to overcome technical issues…So, so many of them were about the credits not being updated, people used to whinge about this but people didn't seem to care so much about if other things weren't updating like the screensaver'* (Interview cpdn-a/2)

Not only does this suggest that the interviewee gained the impression that volunteers were not interested in the screensaver, confounding expectations, but that they were also very interested in one feature of the project that the project team members had previously seen as trivial to the extent that its inclusion was only decided upon at the last minute. Thus, as well as perceiving that their initial expectations of how to retain volunteers was mistaken, they also started to interpret what they observed of the volunteers as meaning the volunteers were primarily interested in, and motivated by, the credit system and participation in the online community.

As well as the experiences described above, the team members interviewed drew on a range of other observations of the volunteers to reinforce their beliefs that competition amongst volunteers based on the statistics systems was the primary motivating factor, even to the extent that some volunteers were willing to engage in activities that were costly and disruptive (to themselves, to the online volunteer community, and even to the project team members themselves) in order to improve their own personal league table positions or the positions of their teams. For example, one interviewee gave the following example:

> **Extract 6.8:** *'That's when things get complicated, when people try to steal members of another team, and it does get quite heated at times, it's one of those things that gets quite difficult to manage, I often say I wish there had never been a credit system to begin with'* (Interview cpdn-a/5)

The interviewee here is referring to instances on the project forums where members of one team actively try to convince members of other teams to defect, leading to arguments on the forum and work for both the forum moderators (and in extreme cases, members of the climateprediction.net team) to attempt to resolve conflict and possibly even take disciplinary action against offending members. In such instances, certain volunteers might be understood as willing to provoke bad feelings in the online community and also to generate more work for others simply in order to improve their own team's league table ranking.

Stories were told in interviews of other instances which were interpreted by project team members as evidence of the overriding importance of the credit system to volunteers. For instance, one individual was discovered to be boosting their credit score by writing a piece of shareware which, unknown to those downloading it, had a hidden version of BOINC embedded in it. Each person who installed this shareware had unwittingly been running climateprediction.net models on their computer, with the credits accruing to the account of the person who had written the shareware. The interviewee who recounted this story did not see this event – and the effort involved in producing the shareware – simply as one person trying to cheat the system but rather as indicative of the high importance of the credit system to the volunteer community in general[40]. Another anecdote which was taken as evidence of the importance of the credit system to volunteers because they were willing to part with money is summarised in the following quotation (again a single example of behaviour is linked to the perceived importance of the credit system by volunteers in general):

> **Extract 6.9:** *'[A credit] doesn't have any monetary value at all, but people were actually trading it on eBay. It's really looked at as a vital factor by volunteers'* (Interview cpdn-a/5)

Thus, it can be seen that the perceptions of the climateprediction.net team regarding how volunteers might be retained by the project shifted after the project launch, as a result of their observations of volunteer behaviour on the project forums, their interactions with various volunteers, and specific events involving the attempts of a small number of volunteers to increase their credit score.

Additionally, shortly after climateprediction.net was launched, a series of articles were published by those involved in SETI@Home and the development of BOINC asserting that the accumulation of credits was a key motivational factor for volunteers (for instance, Anderson 2004). It is likely that the climateprediction.net team were influenced by such reported experiences, for a number of reasons[41]. For instance, SETI@Home was at the time, the most prominent CCP in terms of number of volunteers, and thus their reported experiences of running SETI@Home are always likely to be respected and influential – indeed, as we have already seen, the climateprediction.net team's initial decision to include a credit system has been ascribed to SETI@Home's influence. Furthermore, as climateprediction.net adopted the BOINC software developed by the SETI@Home team in 2004, their contact with SETI@Home team members rose, for example as a result of attending the annual BOINC Workshop. Indeed, traces of the influence of SETI@Home team members surfaced in my interviews with climateprediction.net team members, e.g.:

[40] This is the interviewee who asked that the interview not be recorded (Interview cpdn-a/1).

[41] This experience is furthered echoed by teams involved in setting up and running other large-scale BOINC-based projects, such as the Folding@Home project: when reporting eight years' worth of experiences with this project, Beberg *et al*. (2009) describe the credit system as *'the heart and soul of any distributed computing project'* (p. 5).

Extract 6.10: *'It became apparent when SETI@Home was launched that this was very, very vital, if not extremely essential to any project, if you don't have it people won't do any work for the project at all'* (Interview cpdn-a/2)

Overall, the beliefs of climateprediction.net team members shifted away from viewing volunteers as people who, once they have learnt about the science underpinning the project and the importance of advancing the scientific understanding of climate change, would inevitably give their continued support to the project. Instead, their beliefs shifted towards perceiving volunteers as primarily motivated by enhancing their social status amongst the community mediated through the project forums, coming to the belief that the volunteers constituted this community according to a hierarchy based on credit scores and credit league tables (both team and individual).

In this sense, then, the volunteers have been conceptualized by project team members as having a deficit of disinterestedness and communism or disinterestedness. Rather than participants regarding their contributions as gifts to the community, climateprediction.net team members instead suggest that volunteers attempt to exercise property rights over the work performed by their personal computers: as discussed above, interviewees frequently mentioned that volunteers would complain swiftly and insistently if their credit scores were not being updated, suggesting that they were keen to be assigned the credits to which they felt entitled to by virtue of their contribution to the project.

Boundary work: A market economy of knowledge operating amongst volunteers? From the above discussion, it can easily be seen that the project team members suggest that a market economy of knowledge, violating Mertonian norms of science, underpins the behaviour of volunteers, and the community they form. But they also make clear that they themselves do not engage in such an economy of knowledge, drawing contrasts between themselves and project volunteers. This is made clear in the following two quotations from my interviews with climateprediction.net team members, in which they suggest that they are unable to understand why volunteers might be motivated by credit acquisition, rather than the good of science, thus suggesting differences between themselves and the volunteers in terms of their relationship with science. The first quotation considers the reasons for why volunteers might be motivated by credit scores:

Extract 6.11: *'There's probably a lot of psychological factors that I'm not aware of but we really don't think about it, I don't think I would ever understand it'* (Interview cpdn-a/5)

Note here that not only does the interviewee state that they don't understand this motivation, but that it defies understanding, suggesting that they see this motivation as irrational. This is echoed in the following quote, where another interviewee reflects upon the incident described above in which an individual hid a version of BOINC in a piece of shareware:

Extract 6.12: *'I don't really know actually, I could never imagine myself doing something like that, so I don't understand what's going on in their minds…Some people do it just for pure competition rather than science'* (Interview cpdn-a/3)

At this point, it might be argued that some of the project team members involved in the project launch have contributed to this market economy of knowledge because it was they, after all, who made the decision to implement the credit and team systems. This, in turn, suggests they have violated the Mertonian norms that might underpin a gift economy of knowledge, in the sense that through this, they endowed an individual volunteer's contribution with value not in terms of its meaning to the volunteer (i.e. as a sacrifice on the part of the volunteer), as in a gift economy, but in terms of its value to the project (i.e. the quantity of scientific work performed), as in a market economy of knowledge.

However, although they did indeed provide the credit system, these team members perform discursive moves that locate the place where credits are endowed with value away from themselves and firmly in the domain of the volunteers. In a number of the quotations (both from my interviews, and in publications), the team members stress that when they implemented the credit system, they believed it to be a trivial and somewhat meaningless feature. For instance, as discussed above (and illustrated above by a quotation), the decision to include the credit system was made only very shortly before the project launch, and was very nearly not included at all. This is echoed by Christensen *et al.*, who, in reference to the importance of the credit system in volunteer retention, asserted that *'climateprediction.net was late in appreciating this aspect of volunteer computing projects'* (Christensen *et al*. 2005, p. 4). The project team members, therefore, make it clear that they never intended or expected credits to be seen as valuable in any respect, and that they have value only in the context of the climateprediction.net volunteer community. Thus, in this respect, the credit system is used by the team members to further highlight differences between themselves and the volunteers. This not only completes the shift in terms of the team members' understandings of the volunteers and their motivations to participate, but also completes the discursive construction of a boundary between the team members and the volunteers.

6.4.3 Comparing the approaches of climateprediction.net and Galaxy Zoo to volunteer retention

At this point, I will take a step back from considering climateprediction.net only and instead make comparisons with Galaxy Zoo in order to clarify both elements of my approach to studying the projects and the conclusions I have drawn, in particular, my arguments regarding how the teams in both projects have come to apparently contradictory ideas regarding the role of credit systems/volunteer statistics/league tables in motivating and retaining volunteers. The teams running both projects have sought to recruit and retain a large corpus of volunteers:

the climateprediction.net team appear to have concluded that a credit system is an essential factor in doing so, whilst the league table in Galaxy Zoo was abolished in part because of fears that it was proving off-putting to many volunteers who could never hope to achieve a high ranking.

However, it must be stressed that I have been focussing on the choices as they were made by the respective projects' teams in the particular contexts in which they were made, *as those contexts were perceived by the decision makers at the time*. These beliefs clearly differed across these contexts. Different members of the two projects' teams may have observed volunteers behave in different ways, or been subjected to different influences/pressures/incentives as they interpreted this behaviour. For instance, as reported above, one climateprediction.net interviewee became aware of credits being sold for money on eBay, influencing them to believe that credit accumulation was important for volunteers. By contrast, a Galaxy Zoo subforum set up shortly after the project began asked volunteers what motivated them to participate: volunteer responses suggested to the core team members that volunteers would not miss the league table if abolished.

Another reasons for why apparently contradictory decisions regarding credit systems and league tables might relate to the scientific work being carried out by each project. This is that, at all stages of project design and operation, the climateprediction.net team have been very influenced by the SETI@Home project, because both are similar in that they ask volunteers to donate computing capacity to process data. As I explained above, in 2002 and 2003, a series of articles were published by those involved in SETI@Home asserting that the accumulation of credits was a key motivational factor for volunteers, and some of my climateprediction.net interviewees suggested that they had been influenced in their views of the credit system by reports from SETI@Home. Furthermore, as climateprediction.net adopted the BOINC software developed by the SETI@Home team in 2004, their contact with SETI@Home team members rose, for example as a result of attending the annual BOINC Workshop. By contrast, there is greater distance between Galaxy Zoo team members and SETI@Home team members and thus it might be supposed that they are less likely to be influenced by reports from SETI@Home about the role of the credit system.

It might also be seen that there are differences between what each project's team regards as valuable contributions by members of the project to the scientific work of the project, and these differences may have proven critical in different choices regarding how to credit volunteers. In the case of Galaxy Zoo, fears were expressed that a league table encouraged volunteers to speed up classifications, and the project team were actually keen to actively slow volunteers down so that the data generated would not be too skewed by individual biases and also to increase the chance of volunteers making serendipitous discoveries – something which emerged as an important scientific goal of the project in its own right and also as contributing to the project's original scientific purposes. By contrast, it is believed that climateprediction.net's scientific goals are best pursued by encouraging volunteers to process data as quickly as possible: a climate model is presumed to yield the same results irrespective of the speed at which it has been run, and the project team see little prospect for other contributions (e.g. serendipitous discoveries) by volunteers.

6.5 Conclusion

This Chapter examined a project which has followed a very different approach to Galaxy Zoo in terms of acknowledging the contributions of its volunteers to its scientific work.

By considering the historical, institutional, political and scientific contexts in which the project was launched, as well as considering the personal goals of certain key individuals within such contexts, it could be seen that the project was launched with the expectations that it would both result in scientific work of significance in the field of climate change forecasting, and that it would engage a large mass of members of the public. Initially, it would seem that the project team members conceptualised the task of retaining volunteers, and motivating them to contribute more, in terms of a form of the deficit model that suggested volunteers might be willing to adhere to the Mertonian norms of disinterestedness and communism but lacking the necessary knowledge to adhere to universalism and organized scepticism, but that this expectation was confounded by experiences and the ways in which they interpreted these experiences. This process of social learning was influenced by a number of events, whose interpretation was in turn influenced by the reported experiences by members of the SETI@Home team, as well as by the nature of the scientific work that the volunteers were being asked to do – where quantity matters and does not impact upon quality. This resulted in a shift on the part of the team members to conceptualizing the volunteers as being, in general, motivated by the accumulation of credits and other statistics in order to pursue social prestige in the volunteer community, i.e. a form of the deficit model that regarded volunteers as lacking the norms of disinterestedness and communism.

This has enabled the team members to be able to give an account of how their project had managed to generate significant scientific output and also to retain a large number of volunteers, as well as enabling these team members to discursively construct and reinforce a boundary between themselves and the volunteers. My attention will now turn to what happens on the other side of this boundary.

Chapter 7: Relationships amongst climateprediction.net volunteers

7.1 Introduction

From considering how the experiences of the statistics systems to one particular group (project team members) emerged and changed temporally as they sought to pursue their interests, I will turn, in this Chapter, to considering how the social reality of the credit system varies across volunteers.

The aim of this Chapter is two-fold. One of its aims is to complement Chapter 6, in particular to examine how the social reality of the credit and statistics systems varies across contexts by serving different purposes for different people. The second aim is to extend upon the social science studies already conducted of volunteers in CCPs. As I explained in Chapter 1, subsection 1.1.2, these tend to fall into two categories:

1) Those that conceptualise a project's volunteers as a single mass of people whose behaviour, motivations and so on are treated as statistical variables and which can be captured in sets of descriptive statistics (means, standard deviations etc.) (for instance, Nov *et al*. 2010; Raddick *et al*. 2009). Thus, these studies do not attempt to differentiate between groups of volunteers; and
2) One study which looks at one group of volunteers only (team members who contribute the majority of a project's work) (Holohan & Garg 2005).

I hope to enrich these understandings through differentiating between two groups of volunteers, highlighting differences as well as commonalities between groups. In subsection 7.2.1, I will present the groups I discuss in this Chapter, describing their main characteristics, why they each make a distinctive and critical contribution to the pursuit of climateprediction.net's goals, and how they each differ from other volunteers. I will then deal with each group in turn (subsections 7.2.2 and 7.2.3). By considering contexts both within and outside climateprediction.net in which they exist, I will explore how the credit system is important to them and for what purposes, in particular highlighting how it has a multiplicity of meanings and uses.

The first group are analogous to those considered by Holohan & Garg (2005), and my efforts here are to extend their analysis of this group, in particular, how their behaviour is incentivized by the credit and statistics systems. The second group are those who put themselves forward for alpha-testing of new BOINC-based projects, and have generally had a high level of education and professional experience relating to areas relevant to the project (meteorology, computer programming etc.). Although the credit system does not have a similar effect on this group as the other group, it has nevertheless played a significant, if surprising, role in their experiences with climateprediction.net, and I will attempt to show how they use the statistics systems discursively in order to attempt to differentiate themselves from, and assert authority over, other volunteers.

7.2 The credit system in the online lives of volunteers

In the previous Section, I discussed how the climateprediction.net team members' understandings of the volunteers have changed during the course of the project. In particular, it appears that this involves a belief that volunteers are motivated both to continue participating in, and to increase their contributions to, the project by the accumulation of statistics to enhance social standing in the online volunteer community and thus the statistics systems are the critical factor in the project achieving both of its major goals. Further, they perceive volunteers as a somewhat homogeneous group, at least in terms of their motivations.

The aim of this Section is to consider the role and importance of the credit system in the lives of these volunteers, questioning the extent to which this matches the team members' understandings. I seek to problematize the homogeneity of the volunteers, through considering the contexts (both on- and offline) in which they are embedded, studying how they seek to situate themselves relative to others (both volunteers and team members) and how they make use of the statistics systems to do so.

7.2.1 Heterogeneity of volunteers

Through my analysis, I have identified a number of different groups of volunteers, based on the extent and nature of their participation in climateprediction.net and other CCPs. These groupings were arrived at using various pieces of information available on the participants' profile pages on the climateprediction.net website, such as their Recent Average Credits, team membership, and the other BOINC-based projects in which they are involved. They were further refined by studying their contributions to the project forums, for instance considering which topics they tend to comment on and how they situate themselves in various disputes (for instance, with which other volunteers do they align themselves? Whose side do they generally take? What sort of arguments – technical, social, scientific etc. – do they tend to employ?).

Owing to space constraints, the results presented below focus on two of these groups. One of these (which I have labelled *Super-Crunchers*) aligns well with the group identified by Holohan & Garg (2005), and I will examine the extent to which their conclusions are supported by my observations, largely finding that they are. Holohan & Garg identified this group as being the critical contributors to the scientific goals of BOINC-based project. However, I have also identified another group (which I have labelled *Alpha-Testers*) who also make a critical, albeit different, contribution to the goals of climateprediction.net. Furthermore, my observations show that the statistics systems have served very important purposes for this group of volunteers – again in a different (and unexpected) manner to the role they play for the Super-Crunchers.

I have decided to focus upon these two groups for a number of reasons. One is pragmatic, namely that it was relatively straightforward, in the cases of both groups, to assemble a sizeable corpus of volunteers with whom I was able to both conduct an in-depth interview and to observe extensive online activity (such as in online forums). Another reason is that, in my view, these two groups make the most critical contributions, by some distance, to the project's scientific goals. Finally, I found that the statistics systems played a very significant role in the online lives of both groups and this makes for interesting comparisons.

Before proceeding with my analyses, I will first describe these two groups in more detail.

Super-Crunchers Super-Crunchers form approximately 10 per cent of all active volunteers within climateprediction.net (the figures given here are derived from a random sample of 1000 active climateprediction.net volunteers). Usually active members of a team, this class of volunteers is characterized by the relatively very large quantity of project data that they process on their own computers (their Recent Average Credit score is usually in excess of 1000), often running BOINC on a number of computers simultaneously for many hours each day (some keep their computers running for all 24 hours every day), and they have frequently adapted their computers to increase their processing capacity. Unlike many other volunteers who also display a strong commitment to participating in VCPs (including the Alpha-Testers, and those who may become forum moderators in a particular VCP), the Super-Crunchers tend to restrict themselves to only a handful (usually between one and three) of VCPs, which allows them to accumulate a high number of credits in specific projects (rather than spreading their computational resources, and hence credit scores, more thinly over a larger number of projects).

As a result of this, the Super-Crunchers participating in climateprediction.net can be seen to make a particularly valuable contribution to the project's goal of producing publishable scientific results: approximately 60% of all credit awarded by climateprediction.net has been awarded to the Super-Crunchers. Indeed, within this class, there is a great deal of variation in the level of work done amongst this group, with over 10 per cent of credits awarded to just 0.2 per cent of active volunteers.

Alpha-Testers The other class of volunteers considered here is called the Alpha-Testers, so-called because it comprises those who volunteer for or are recruited by VCPs that are at the alpha stage of development. They comprise a very small group (approximately 5% of active volunteers). Many well-established VCPs also retain a group of such volunteers to be able to test new software or new types of work units (for instance, in the case of climateprediction.net, the HADam3 seasonal variation sub-project). The Alpha-Testers in climateprediction.net are often registered with 15, 20, 30 or even more other BOINC-based projects.

This group can be seen to have played a critical role in contributing to the project's goals in the sense that they were vital for providing feedback when the project was at an alpha-testing phase, or when it was testing new features (such as new subprojects).

Other volunteers Other volunteers (comprising over 85% of active volunteers) form a diverse group, although strikingly few of them have direct contact with scientific institutions outside of CCPs. Together, they account for a minority of credit score (approximately three-quarters of these volunteers have a Recent Average Credit score of less than 100), and the majority participate in 0-2 BOINC-based projects other than climateprediction.net. It was very difficult to make contact with, and study, these volunteers. They tend to be infrequent visitors and contributors to the online forums and few answered my questionnaire (and were thus not recruited into my study). Furthermore, it is very difficult to build up a sound understanding of individual volunteers who post seldom (if at all), or for a very short period of time only, on the forums[42].

7.2.2 The Super-Crunchers: pursuing social status through the accumulation of credits

The Super-Crunchers enjoy the prestige that high credit scores bring within the online community of project volunteers (both in climateprediction.net's own forums, and the forums of the teams they join). Some Super-Crunchers work to achieve prestige on their own, whilst others operate as part of a team and pursue a higher status through their team's success. For ease of reference, I shall refer to the former as *Solo Crunchers* and the latter as *Team Crunchers*[43]. Solo Crunchers pursue individual glory through achieving a high position in the project's league tables and the reporting of their own credit scores and other achievements (such as the number of climate models they have run to completion) in the project forums for which they receive praise from a broad cross-section of other volunteers. To complement this, a Facebook application was launched in summer 2009, which enabled credit milestones to be displayed on a volunteer's Facebook profile. Furthermore, the Solo Crunchers often display their credit scores in climateprediction.net (and other BOINC-based projects) in the signature they attach to forum posts (see Figure 7.1). I did not get the opportunity to speak with any of these Solo Crunchers in-depth as they are

[42] There were some notable exceptions to what I have written here, including a number of forum moderators.

[43] In reality, there is not so much of a clear distinction between the two as many of those I classified as Solo Crunchers are part of a team. However, in the cases of these volunteers, their team membership is incidental, which is in stark contrast to Team Crunchers, whose individual identities often come second to their team identity (for instance, their behaviour across various BOINC-based projects is directed towards team goals rather than individual goals).

relatively very few in number, however it was clear that they enjoyed receiving the praise of other volunteers for their attainments.

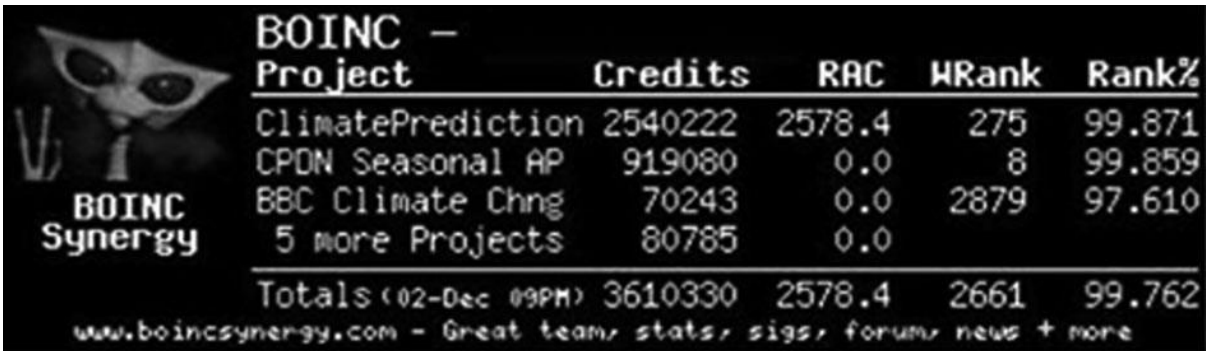


Figure 7.1: A climateprediction.net volunteer's signature. *This gives the volunteer's current total credit score within the main climateprediction.net project and two climateprediction.net subprojects, the RAC and the rank of the volunteer in comparison with other volunteers both as an absolute figure and as a percentile.*

Instead, I shall focus on the Team Crunchers, in particular the example of one such team, called the *Dutch Power Cows* (DPCs). This is a team of volunteers who mainly reside in the Netherlands. As of 31 October 2011, there are approximately 750 active members, active across a broad range of BOINC-based projects. As of this date, the team is ranked first on the Malaria Control project team league table, second in the Leiden Classical project table, in the top 10 of three other projects (including Rosetta@Home, one of the very largest BOINC-based projects) and in the top 20 of a further eight projects (including the very large World Community Grid project)[44]. Although they are ranked in a relatively lowly 30th position on the climateprediction.net league table (as of 31 October 2011) with only a handful of members actively participating in the project, the team was more active in the project during 2008 and 2009, when I was conducting my fieldwork, and were regularly ranked in the top 10 teams.

The DPC team has existed since the early days of CCPs. It has its own website[45], with a variety of features, including and a forum. Team members are able to donate their computer time to whichever BOINC-based project(s) they choose. Those DPC members I spoke to suggested that their involvement in a team was the key reason they continued to participate in BOINC-based projects, for instance:

> **Extract 7.1:** *'I'm a member of the Dutch Power Cows, the largest Dutch team. It brings that extra something to projects that make participating in them fun'* (Interview cpdn-c/3)

In particular, it is the competitive element of the credit and team system that keeps such volunteers engaged in these projects. As I shall discuss below, this competitive element operates on two levels: one is at the level of a

[44] http://boincstats.com (accessed 9 December 2011)
[45] www.dutchpowercows.org (accessed 9 December 2011)

project where members of the team come together to boost their team's league table position, and the other is rivalry between different members of the team as they attempt to improve their relative positions.

At the level of BOINC-based projects, the team attempts to attain a high league table position in as many projects as possible. In order to understand how this motivates team members to participate in particular projects, I will explain the *stampede*. A stampede is defined thus on the team's website[46]:

> *'A stampede is an organized, temporary joining of our strength on a project. During a stampede, which lasts one month, virtual teams formed within the DPC that anyone can join. The internal competition increases our output for that project to unprecedented heights'*

Stampedes are common to a many teams. In a stampede, a particular project is targeted, and all team members are invited to join in and donate their CPU time to that project for a whole month in order to boost the team's position on that project's league table. A quotation from an interview with a DPC member explains further:

> **Extract 7.2:** *'Since we are a big team (ranked 10th overall projects in the world) we take part in a lot of diverse projects but if we tend to lose a good position or only need a bit extra to gain one the members all get together and switch over computing power from all projects to that one that needs it. And being a very stats driven team, not having stats for a few days (not once but a lot of times) has led to people leaving climateprediction.net and joining another project'* (Interview cpdn-c/2)

In other words, the primary reason for choosing a particular project for a stampede is not whether the project offers the opportunity for contributing to science, but rather to compete with other teams. The quotation also clearly links this to the stability and reliability of the project's credit system: if it breaks down or does not update scores regularly, then team members may no longer be motivated to continue participating in that particular project and are liable to switch to another project. This is used to explain why a number of DPC members left the climateprediction.net project (and why there have not been any Stampedes targeting climateprediction.net since 2009) – it is not due to DTC members questioning their support for climate science or whether their participation in the project is genuinely contributing to the project's scientific goals, but because the project's credit system did not satisfy their urges to compete against other teams.

At the level of competition within the team, the credit system of a particular project can also be seen to be important for retaining individual team members. Within the DPC team, members can form subteams which compete against each other, and league tables comparing subteam credit scores for each major BOINC-based project are posted regularly on the DPC forums. This also motivates individual members to continue to participate in a particular project, as well as encourage them to recruit others to join their subteam and thus to participate in that project, as the following quotation from one of my interviews explains:

[46] Note that this is a translation from the original Dutch on the website.

> **Extract 7.3:** *'I am involved in a community battle for the most CPU power so that drives us to the max. My friends and parents all donate their spare CPU time to climateprediction.net, and to my account, so I'll have a high output and we've reached a nice 4th position in our local ranking on DPC. My friends all love the stats, especially when we are high up'* (Interview cpdn-c/4)

Thus, it can be seen here that in the case of a project such as climateprediction.net, it is the competitive element engendered by the credit system and team system – rather than the potential to contribute to the project's scientific goals – that has engaged some Team Crunchers in the Dutch Power Cows in the past, both in terms of the team competing the project's league tables and also internal competition amongst DPC members. This accords to some extent with the findings of Holohan & Garg (2005). They argued that the large credit scores accumulated in BOINC-based projects by volunteers involved in teams resulted from *'co-opetition'*, namely that competition took place within a BOINC-based project amongst teams, and that individuals within a team would cooperate in order to improve their team's performance. Such cooperation would involve posting messages in the forums in the team's website encouraging or urging other team members to improve their performance, offering technical support and advice to other team members regarding the running of BOINC work units, and getting team members to target a particular project at a particular time in order to improve the team's ranking (analogous to the Stampede discussed above).

My findings support this to some extent, suggesting that Team Crunchers are indeed primarily motivated by improving their team's performance in league tables. They may well be interested in contributing to science, and indeed may restrict the projects they will get involved with according to their personal assessments of projects' scientific value, however they are willing to switch their focus from one project to another if the team's interests demand it (e.g. for a stampede). However, my findings also suggest that these Team Crunchers do not necessarily totally subsume their personal identities to their team identities, and that competition can be found within teams. Indeed, it suggests a form of *nested co-opetition* in the sense that, within a team, sub-teams can form which compete with each other and within which, members can cooperate to improve their sub-team's position relative to others: all these factors can contribute, ultimately, to increasing the amount of computer processing capacity donated to BOINC-based projects.

However, Super-Crunchers are only one group, and my studies of these volunteers have revealed other groups of volunteers for whom, despite not being motivated by the credit and statistics systems, these systems have nevertheless played an important role in their online lives in climateprediction.net.

7.2.3 Alpha-Testers: Credits and statistics as a tool in boundary work

The vast majority of climateprediction.net volunteers participate in up to a handful of other BOINC projects, and these projects tend to be well-established, having already passed through alpha and beta testing phases. By contrast, the Alpha-Testers are often involved with 15, 20 or even more than 30 BOINC projects, many of which are in an alpha or beta testing phase, or they are regularly involved in the alpha testing phases of new features of climateprediction.net. Such individuals are very valuable to BOINC projects in a number of ways.

The Alpha-Testers often come from technical and scientific backgrounds, involving fields relevant to climate science (such as meteorology, or environmental science). They usually state that they have degrees in scientific disciplines, and careers that involved extensive computer programming or hardware development. This is exemplified in the following quotation from an interviews with an Alpha-Tester:

> **Extract 7.4:** *'As a former US Air Force Weather Forecaster, I've long had an interest in weather and related topics -- including climate. My association with computers began in 1964, with mainframes. In the first half of my USAF career, I was involved in weather reconnaissance before satellites, and a few years of that involved flying into Typhoons -- an exciting and very interesting way to earn a dollar! Most of the second half of the career was involved with mainframe computers at the USAF Global Weather Central'* (Interview cpdn-b/3)

It certainly seems that the Alpha-Testers are not motivated by the acquisition of credits or the prestige accrued from other statistics (such as the number of climate models completed) – after all, their involvement in CCPs tends to be spread over a large number of projects, rather than specialising in a smaller number, and furthermore, credits at the alpha and beta stages of development are rather unstable (as a project's developers might modify the policies governing the awarding of credit as they refine the project for public launch). Indeed, in my interviews with the Alpha-Testers, they consistently sought not only to say that they were personally uninterested in project statistics but to actively distance themselves from this, and from people who might be motivated by this. One Alpha-Tester stated: *'I find the chasing of meaningless statistics utterly offensive'* (Interview cpdn-b/3) (**Extract 7.5**), whilst another went into greater depth about their apparent contempt for those who are supposedly motivated by improving their statistics:

> **Extract 7.6:** *'For a short time, I joined the Pacific Northwest team but dropped out when it became obvious that members' interest was meaningless credits. I received an angry email from the team founder because…my credits went away when I left the team. Credit whores, as another [forum] moderator calls them, irritate me. Meaningless credits. Superficiality. Kid stuff. It's about the science, not silliness!'* (Interview cpdn-b/1)

Here, it might be seen that, whilst they assert they are not themselves motivated by statistics, they use the existence of the credit system and other statistics to situate themselves discursively relative to other volunteers. For instance in the quotation above, other volunteers are cast by the Alpha-Tester as being motivated by credit acquisition and are thus more interested in superficial *'kid stuff'* rather than in contributing to science. As I stated above, Alpha-

Testers tend to have an educational and/or professional background in fields related to the scientific or technical aspects underpinning climateprediction.net, and thus might be keen to construct and maintain a boundary between themselves and other volunteers. In Chapter 6, subsection 6.4.2, I discussed how the manner in which the project team members cast volunteers as motivated by credit acquisition has discursively constructed a boundary between them and the volunteers. It might be the case that, in the context of the volunteer community itself (i.e. in the online forums), there are few ways for Alpha-Testers to assert a social position based on their own technoscientific background and the credit system, and other project statistics, afford opportunities to engage in boundary work of their own.

To see an example of this, I will present data from one thread in this subsection, which will be referred to as *Thread F*, in the climateprediction.net forums, occasionally bringing posts in other forum threads to clarify or develop points. Thread F was chosen as a focus primarily because it became a site for a dispute in which some Alpha-Testers played a prominent role in which they revealed a number of their attitudes about the roles which others could legitimately play in the climateprediction.net project.

This thread originated as a request for help from a project volunteer (*Rob*) whose HadCM3 model persistently crashed at a relatively late stage in its running. A number of volunteers offer advice on how the volunteer should proceed[47]. A consensus is quickly reached that the model was unsalvageable, and that the volunteer should move on to another work unit. Rob expresses disappointment, as it was the first model he had not been able to complete. Consolation is offered by two forum moderators (*Jim* and *Eileen*) that results from the crashed model would still be of value to the project. Indeed, Jim states that figure reported by climateprediction.net on its website as the total number of models completed by volunteers actually includes all models that have reached 2050, and not 2080, in its reported count of the number of models completed by project volunteers. This provokes a response from two Alpha-Testers (*Jet* and *SG*), both of whom say that it was wrong to define such models as completed when teams only count models which had reached 2080 as complete, and that there is a real need for consistency and lack of ambiguity about the measurement of individual, team and project statistics. It is the conduct of Jet and SG in this thread which shall be dealt with below, looking at how they attempt to define who does and does not have legitimate authority to behave in particular ways, assign tasks to others, and thereby attempt to construct a social hierarchy.

[47] Note that *Rob* is a pseudonym: all the named volunteers here have similarly been assigned pseudonyms.

Alpha-Testers define their legitimate authority In the forum thread considered here, Jet and SG participate in a particular, and limited, way: although the thread was started ostensibly to solicit advice on an issue related to the day-to-day running of the BOINC on a participant's computer, they refrain totally from offering any such advice (or subsequently consolation) to Rob, and participate only in the discussion about the measurement of statistics, where they strongly advocate the importance of strict and universal rules relating to the measurement of statistics.

First, they act to exclude from the topics they seek to present themselves as having legitimate authority to talk about the day-to-day running of BOINC, and climateprediction.net models, on volunteers' computers, even though they could be expected to be very familiar with running the software as they have amassed a high number of credits, and are long-standing volunteers. In Thread F, they do not offer any advice. Indeed, Alpha-Testers seldom offer such advice on the forums, and in the instances they do, they actually present their advice in such a way that reaffirms the notion that they lack the legitimate authority to do so in general. The following post (from another thread) exemplifies their approach:

> *'I have 2 Sulphur wu's [work units] at the moment...In general, if you have 2 wu's...I'd leave them to run. No need to abort them.'* (Post L/005)

Whenever an Alpha-Tester offer advice in the climateprediction.net forum, they cite an instance of personal experience and suggest it is very similar to the predicament of the participant to whom they are giving advice by suggesting common elements, thereby justifying their authority to give the particular piece of advice; by contrast, many threads by non-Alpha-Testers simply assert advice. For instance, advice given by non- Alpha-Tester Jim to Rob begins: *'The model may have reached as far as it is going to go'* – no personal experience cited here – and then offers advice: *'It is generally worth restoring from a backup'* (Post F/003). Such a lack of justification can be understood as an implicit, general claim to authority relating to the giving of on the part of some non-Alpha-Testers; by contrast, the practice of Alpha-Testers of always citing a personal experience lays no claim to authority for giving advice beyond the specific instance where they do.

Instead, they seek to cast themselves as possessing legitimate authority to discuss and to seek to influence climateprediction.net's policies and organization of volunteers. Firstly, they are very vocal in Thread F regarding the measurement of volunteers' statistics, arguing that consistent, well-communicated rules are critical to ensure that volunteers' behave in the best interests of the project. For instance, SG writes:

> *'If climateprediction.net report anything past 2050 as complete, this leaves us singing from different hymn sheets...If the rules of the game are to change, they need to be changed for all teams'* (Post F/008)

And Jet:

'Let's get this crystal clear right now and not leave any misunderstandings, now or in the future...it would be dreadful to find people would 'tactically' prematurely abort a model at 2050 just to show how great their team is, or whatever.' (Post F/027)

Jet's use of the imperative in the latter post is suggestive of attempts on his part to define what he considers appropriate behaviour of others, legitimate ways in which others can and should act, and how he, and other Alpha-Testers, does so is considered later. However, it must first be noted that in so doing, the Alpha-Testers are claiming legitimate authority to define the roles and relative positions of project volunteers and assign tasks to them. In other words, they claim legitimate authority to advocate and enforce their views on the social organization of climateprediction.net.

Alpha-Testers define the roles of others Here, I will focus particularly on the interactions in Thread F of Jet with the previously-mentioned Eileen, Jim, and Rob, as well as another project volunteer (*Sheila*). Both Eileen and Jim are long-standing forum moderators. Sheila is the leader of a relatively large team, and posts a high number of messages relating to teams, welcoming new recruits to the team, and congratulating team members on reaching particular thresholds for total credits acquired or the count of Models Completed. Rob has no official status apart in the climateprediction.net project apart from that of project volunteer.

First, the case of Rob shall be considered. In addition to writing the initial post of the thread soliciting advice, he goes on to make frequent posts relating to how a completed model should be defined, but it is clear that Jet does not regard these as legitimate contributions to the debate, and this can be seen by comparing Jet's responses to Rob's posts to his responses to Jim's. Both make posts almost simultaneously, with both apparently arguing along similar lines, seeking to resolve the controversy about what counts as a completed model by appealing to the good of science. They both argue that, as models approach 2080 (and pass 2050), they are unlikely to provide accurate predictions of what will actually happen, and hence results for the 2050-2080 time period are of reduced worth to the scientists and it is therefore legitimate for them to consider models which have reached 2050 as 'complete'. Some hours later, Jet posts a response, which is oriented solely towards the post by the moderator, Jim. This can be seen firstly because the post is specified as being in response to that of Jim's and also because Jet directly quotes a phrase from Jim's thread *(' "please think about the physics and science of it."'* (Post F/020)).

This was even in spite of Rob's post actually being addressed to Jet (*'Hi [Jet], my fellow Team England (BOINC) team mate'* (Post F/015)). This is typical of Jet's behaviour towards others in this thread: respond to those posts about the measurement of statistics made only by moderators or Sheila, a team leader, and ignore the many made by others. Every post made by Jet on the issue prior to its apparent resolution (with one key exception, discussed in the next paragraph) either states that it is *'in response to Message ID [number of a post made by a moderator or Sheila]'* and contains direct quotes from the posts of moderators, or Sheila, only, creating the strong impression that Jet regards the contributions to the debate of those who don't hold an official post as not having any value.

This can be further understood as also forming part of a strategy on the part of Jet to assign the task to these three individuals of communicating unambiguously and explicitly to the mass of project volunteers regarding which models count as 'complete'. The first part of the strategy involves the construction by Jet of a division between these three individuals, casting them as being able to understand how to best serve the good of science, and the broad mass of project volunteers, who are portrayed as lacking this understanding, instead needing to have their behaviour directed and regulated by the system of attribution of credits and other statistics. For example, when addressing Sheila, Jet states: *'Science is fair and obvious to you'*, but then goes on*: 'it's so obvious that winning on credit is more important that (sic) the climate science and...It's quite clear that* "completed models" *is a top bragging item for many'* (Post F/030).

The next part of the strategy involves Jet suggesting that any confusion among the volunteers about when a model can be counted as completed is damaging to the interests of the project, because it might induce volunteers to *'*"tactically" *prematurely abort a model at 2050 just to show how great their team is'* (Post F/027), and that the fault of such confusion lies with people such as the moderators or team leaders. For instance, Jet says to Eileen: *'The fact is you have given an impression that models at 2050+ are complete. I expect some are confused about this'* (Post F/013). By accusing them of being the source of the confusion amongst the volunteers, Jet is also portraying them as already possessing authority to influence and direct the project volunteers. In turn, Jet is claiming authority to influence these three individuals regarding the way in which they control the project volunteers, and attempts to do so to get them to unambiguously make it clear to the project volunteers that only models which get to 2080 should be considered complete.

Briefly, then, the behaviour of various Alpha-Testers in Thread F, and elsewhere, suggests that they believe in: the critical importance of consistent and clear measurement of volunteers' statistics (count of models completed, credits etc.); strict demarcations of when volunteers can and cannot act with legitimate authority; and, embedded within this, a social hierarchy where Alpha-Testers are able to discuss and influence climateprediction.net rules and policies, and designate moderators and team leaders as being responsible, and possessing the authority, to enforce these rules amongst project volunteers.

Theoretical models: Legitimate authority At this stage, a theoretical framework will be introduced which might help to explain the data presented above will be considered. This is the model of *'legal authority'* proposed by Max Weber, one of *'three pure types of legitimate domination'* (Weber 1947, p. 31). By a *'pure type of legitimate domination',* Weber meant the existence of a valid social order, valid in the sense that individuals will orient their behaviour towards the norms of the social order. However, an important point to bear in mind below is that in *Economy and Society*, Max Weber argues that the existence of a legitimate social order does not necessarily imply that all actions will obey the social order. In the case of an act which transgresses the social order, the order is nevertheless recognised as legitimate if, for instance, the transgressor feels compelled, publicly and explicitly, to justify their behaviour.

The main features of a Weberian rational-legal social order are (Weber 1946, 1947):

- Clearly-defined rules to regulate behaviour, enforced by a hierarchical authority structure, with authority delegated down the structure;
- The existence of an impersonal order behind these rules which even the individual(s) in authority must obey;
- Clearly-defined *spheres of competence* for each individual. An individual has the authority to perform any action which falls within their sphere, and are trusted to do so competently, and it is considered inappropriate for them to act outside of their sphere; and
- The existence of a reward system, with reward directly related to the quantity of work done towards the ultimate goals of the social order, as an incentive for individuals, and hence can be seen as foundational for the functioning of the social order.

The Alpha-Testers studied above appear to support a social order whose organization is founded upon rules relating to how an individual's models are counted as completed and credits are attributed. As measurements of work done, and the kudos gained by volunteers when they display their statistics on the online forums, they can be seen as functioning in the same way as a reward system in a Weberian legal social order, with individuals (and teams) being able to display their statistics to other volunteers in the forums to gain kudos, and hence can be seen as a way in which climateprediction.net can direct the behaviour of its volunteers through giving them incentives to behave in particular ways.

Furthermore, Jet's (and other Alpha-Testers') behaviour, in the forums appear to be akin to seeking to define spheres of competence both for themselves and for others, and furthermore, spheres which are organised into a social hierarchy. Alpha-Testers regard themselves as legitimately competent to discuss rules regarding how statistics are measured and, given that they believe these form the basis of attempts to direct and govern the behaviour of project volunteers, it can be seen that the organization and policies of climateprediction.net may be supposed to fall within their spheres of competence; by contrast, giving advice on the day-to-day running of cliamteprediction.net models and software on volunteers' computers falls outside their spheres. Weber stated that the recognition of the legitimacy of a social order by an individual does not always means they will obey the social order, but that they will act in a way to acknowledge that they realise their transgression and to legitimate the particular instance when they transgress. In a similar vein, although Alpha-Testers do occasionally offer advice technical advice, they always relate it to a particular experience that they have had.

Boundaries and hierarchies In the example above, we can see that the Alpha-Testers above used a dispute regarding a policy by which one of the statistics systems operate to try and construct a boundary between themselves and the other volunteers. It was argued that these Alpha-Testers appear to advocate, and indeed enforce, a social organization founded upon these rules matching very closely to a Weberian rational-legal system. One of the key features of this system is the assignation of spheres of competence, and the Alpha-Testers acted in a way which seemed to reveal strong beliefs that particular volunteers possess legitimate authority, and responsibility, to act in particular ways and not in other ways. Of particular note was the way in which Alpha-Testers sought to enforce a social hierarchy, aligning themselves with those who set up and maintain climateprediction.net and designating forum moderators and team leaders as having the responsibility and legitimate authority to control and direct the volunteers. Furthermore, the Alpha-Testers argued clear, consistent, and well-communicated rules regarding the measurement of project statistics are necessary in order to direct the behaviour of project volunteers.

Such a system is a structure for organising a mass of people in order to maximise output based on the assumption that these people are able to judge for themselves not how to behave for the good of the whole but how they can pursue their own self-interest. and thus must have their behaviour directed through a system of personal rewards. We can see here that the Alpha-Testers are suggesting that the other volunteers do not understand how (or are not willing) to behave in a way that subsumes their own interests for the good of climateprediction.net, and instead are interested in attaining status through improving their statistics. This is akin to the conceptualisation of volunteers implicit in the discourse of the climateprediction.net scientists (see Chapter 6, subsection 6.4.2). At the same time, the Alpha-Testers are placing themselves on the other side of this boundary: by setting down to others how they believe the project should be organised, and attempting to define the roles and tasks of others in this, they are implying that they, themselves, do understand how the volunteers should behave in order to improve the project's pursuit of its scientific goals. Thus, we can see that these Alpha-Testers – who have come to the project from computer programming and environmental science backgrounds – use the statistics systems to attempt to secure a status relative to the other volunteers that recognises their expertise in the technical and scientific issues of the project.

7.3 Conclusion

In the previous Chapter, I argued that it was used by team members to interpret why volunteers were motivated to continue participating in the project, to provide an account to others of why this is the case, and to discursively reinforce a boundary between themselves and the volunteers. Further, I argued that these uses only emerged over time. In this Chapter, I have argued that the credit and statistics systems have been used by one group of volunteers (the Super-Crunchers) as the basis of a system of competition[48], whilst another (the Alpha-Testers) has employed these systems in a series of discursive moves to differentiate themselves from all the other volunteers and to align themselves with the project team members[49].

[48] In future work on developing improved understandings of Super-Crunchers, it might be useful to draw upon studies carried out on the culture of those who engage in online gaming (for instance, see Morris 2002, White 2006).

In common with the credit and team systems in BOINC-based projects, online gaming can bring people together to compete over the Internet. Furthermore, many such games involve the formation of *clans* or *guilds*, each with their own core identities, and in which players come together to cooperate and engage in competition with other clans or guilds (Williams *et al.* 2006). This can easily be seen as potentially analogous to the team system of BOINC.

At first sight, it may appear that the scope for comparison between Super-Crunchers and members of online gaming culture is quite limited – after all, the core task of a climateprediction.net volunteer (to download BOINC and run climate models in the background on their personal computer) may not seem particularly challenging or satisfying when compared to the tasks undertaken in gaming settings such as *World of Warcraft* (for instance, Whippey 2011). However, success for Super-Crunchers and teams requires overcoming a number of challenging obstacles, which can therefore prove very satisfying to those involved.

For instance, there are technical challenges involved in building or improving computers to improve the rate at which a volunteer processes climate models. There are also social challenges involved when a team seeks to recruit other project volunteers, with teams responding to these challenges by building websites to foster a sense of camaraderie amongst members, actively posting on project forums, or seeking to develop a team identity which is attractive to potential team members. Furthermore, as mentioned in subsection 7.2.2, team members often try to recruit friends or family members to participate in a CCP and to join their team (or subteam), and thus the challenge of tapping offline social networks may (if successfully accomplished) provide another element of satisfaction for volunteers. Hence, it can be seen that the challenge of improving credit scores and league table positions might be understood as a multifaceted game.

[49] In future work on Alpha-Testers, it may prove fruitful to draw on the body of ethnographic work studying hacker culture. In this context, the word 'hacker' is not used in the pejorative sense that it is used in popular media to describe somebody who is illegally seeking to gain access to, and possibly tamper with, computer systems. Instead, it is applied to an individual who is very skilled in computer programming and uses these skills to improve existing computer code or build new code. The work of hackers is often collaborative, and gave rise to a community with a distinctive ethos (Coleman & Golub 2008; Levy 1984). Out of hacker culture grew the Open Source movement, which stresses sharing of code and openness, amongst other values (Lindenschmidt 2000; Raymond 1999).

Alpha-Testers work with a CCP's team of software engineers to improve the technical functioning of the project, and thus it is unsurprising to find a number of similarities with the characteristics of hacker culture identified in the literature that studies it and the behaviour of Alpha-Testers documented in Chapter 7. For instance, in his history of hacker culture, Levy (1984) describes the disdain that the very earliest hackers, at MIT in the late 1950s and early 1960s, had for other students. The hackers eschewed to a large extent the study required to pass formal classes and examinations (and thus the accumulation of grades), instead spending their time and efforts on computer programming (which they perceived as truly important and interesting). Similarly, we saw how the Alpha-Testers take great care to present themselves as disdainful of other volunteers' eagerness to accumulate credit scores, which they regard as trivial.

At the start of this Chapter, I set out three aims. One was to extend the social science analyses already carried out of volunteers, and I have differentiated groups of volunteers, and identified and discussed in detail one group that had not previously been covered (the Alpha-Testers). Another aim was to demonstrate how the social learning of the team members has led to a partial perspective: as I argued in the previous Chapter, their accounts have presented volunteers as a homogeneous group, whilst the analysis in this Chapter has exposed (to a certain extent) their heterogeneity.

Finally, I hoped to further demonstrate that the social reality of the credit and statistics systems can vary across contexts. As I explained in Chapter 4, subsection 4.1.1, Mol argues that things can have a multiplicity of identities across a range of contexts. Here, I have highlighted that one element impacting upon the multiplicity of the credit and statistics systems relates to the different uses that different people put them to in climateprediction.net as they pursue their own, different, interests. However, it must also be stressed that these different contexts do not exist independently of each other, nor do the ways in which the credit and statistics systems are used. For instance, the Alpha-Testers and the team members are able to use these systems in the ways that they do because there are many volunteers who do display behaviour that suggests they are primarily motivated by the acquisition of credits: this has enabled the Alpha-Testers and the team members to cite such behaviour in support of their suggestions that all volunteers are motivated by statistics. Furthermore, in the forum thread analysed in subsection 7.2.3, the Alpha-Testers were only able to use the statistics systems in the manner they did because one of the Super-Crunchers had, out of concern for their own statistics scores, raised the issue of whether they could count a model run past 2050 as complete. By contrast, the Super-Crunchers are able to indulge their competitive natures precisely because the team members first implemented a credit system with teams and league tables, and had subsequently ensured that the policies for awarding credit have been consistent over time, thus providing a firm foundation for competition. Thus, the multiplicity of purposes served by the credit and statistics systems, is brought about by the differences in contexts between different project participants, but also heavily influenced by relationships between these contexts.

As a corollary to their disdain for degrees and formal certification, Levy (1984) found that hackers did not believe that individuals should be judged on criteria such as degrees or official positions. Instead, judgements should be based on their hacking contributions only. Again, this points to potential similarities with the attitudes and behaviour of Alpha-Testers. For example, it was shown in subsection 7.2.3 that they appear to believe they can speak with the same authority as that of the climateprediction.net team of scientists and software engineers. Here, then, it might be the case that the Alpha-Testers perceive that formal position titles do not matter; their authority instead resting on their contributions to the technical side of the project.

A third, and final, similarity that I will highlight here is *flaming*. Flaming refers to a practice in which individuals deliberately provoke (sometimes acrimonious) arguments with the aim of improving technical or organisational procedures (Leach et al. 2009). The forum thread analysed in subsection 7.2.3 contained a hotly-contested debate, which began when an Alpha-Tester quite pointedly criticised the practices of other volunteers in terms of how to count a model as 'completed'. Indeed, it is quite usual across the climateprediction.net forums for arguments to be provoked by the posting of an Alpha-Tester.

8 Conclusions and Future Work

In this thesis, I have attempted to answer the following Research Questions:

> **Research Question 1:** How do scientists and software engineers who are about to embark upon Citizen Cyberscience Projects (CCPs) intend to establish, structure, and maintain relationships with members of the public?
>
> **Research Question 2:** How do those with some experience of setting-up and administering CCPs seek to establish, structure, and maintain relationships with project volunteers?
>
> **Research Question 3:** How do project volunteers seek to establish, structure and maintain relationships with other project volunteers?

In order to do this, I employed a bottom-up, qualitative ethnographic approach to studying the communities that have been mediated through two CCPs, climateprediction.net and Galaxy Zoo. By situating these projects, and the individuals involved, in their broader contexts (institutional, historical, political, scientific), it has been possible to understand how these projects have come into being, how they have developed as a result of choices made by key decision makers, and, in particular, how relationships between various project personnel have been structured, mediated, contested, maintained, broken down, and so on.

In this Conclusion, I will start by presenting my main empirical findings (from Chapters 4-7). I will then discuss these in light of the academic literature presented in Chapter regarding how scientists and laypeople interact with each other, and how laypeople interact amongst themselves, around issues of technoscience (Section 8.2). Then, I will consider how my findings extend existing social studies of CCPs. In particular, I will discuss how they might prove useful for scientists and software engineers involved in setting-up and running a CCP, and, to this end, I will provide some recommendations for practice (Section 8.3). Finally, I will consider how the work in this thesis could be extended in the future, in ways that might prove of value to those interested in improving understandings of how CCPs work in order to improve their running in the future, and scholars in the field of Science and Technology Studies (Section 8.4).

8.1 Key empirical findings

8.1.1 Chapters 4 & 5: Findings from the Galaxy Zoo case study

The findings I presented in Chapters 4 and 5 were above all concerned with the relationships between scientists and volunteers in the Galaxy Zoo project and the way in which these relationships are constituted, (i.e. these findings relate to RQs 1 and 2). In particular, I was interested in an ethos that has apparently emerged in the project regarding how the project scientists (both core team members and scientists enrolled to should treat the volunteers, and the conditions under which objects and knowledge are exchanged amongst scientists and volunteers. My core argument was that this ethos did not, however exist prior to, and independently of, the decisions that the core team had to make during the setting up and running of Galaxy Zoo. Rather, I attempted to show how this ethos has emerged during the course of the project, how it has been shaped by (and in turn shapes) the interests of various project members, and, furthermore, how the technical and scientific aspects of the project have driven (and been impacted by) this ethos and these interests.

This argument was split across Chapters 4 and 5 as follows: in Chapter 4, I introduced the project and considered the contexts into which it has emerged, before turning to the interests of the team involved in the setting-up and running of Galaxy Zoo, how they have pursued these interests to date through Galaxy Zoo, how these interests have changed over time as they have discovered (and pursued) previously-unanticipated opportunities afforded by the project, how this has impacted on the way the project has been developed, and the potential threats to these interests that these team members see looming on the horizon. In Chapter 5, I then described the key elements of the ethos regarding how Galaxy Zoo scientists should treat volunteers. By focussing on particular features and policies of the project, and their development during the course of the project, and drawing on the conclusions of Chapter 4, I discussed how this ethos has emerged as a result of choices which the project's core team believed would better enable themselves to pursue their emergent interests. I also argued how the features and policies considered – and the discourse surrounding them – constitute a particular form of knowledge economy, namely a gift economy, that governs the conditions of exchange amongst volunteers and scientists and that this suggests the operation of Mertonian norms.

However, it should be stressed that this does not mean that the project scientists operate unproblematically according to these norms: one of the critical elements of these norms is the idea that the scientist completely sets aside any self-interest in the service of science, but it was seen that the emergence and stabilisation of the Galaxy Zoo ethos itself has been driven to a large extent by the project scientists making choices in accordance with their own interests, that this ethos is indeterminate and subject to constant negotiation and possible change, and thus may take a very different form (and thus not necessarily suggest Mertonian norms) in the future.

8.1.2 Chapter 6: Findings regarding climateprediction.net scientists and software engineers, and comparisons with Galaxy Zoo

Chapters 6 presented some findings from my study of climateprediction.net, focussing in particular on the project team members (scientists and software engineers) and their dealings with volunteers (thus, this Chapter can be seen as further answering RQs 1 and 2). First, I situated the project in its historical, scientific, and social contexts, in order to understand how the project emerged and how its goals are oriented towards these contexts. I then explained the major features of the project, and how the project has developed over time. This set the scene for charting the team members' changing perceptions of the volunteers, in particular how, as the project unfolded, these team members have developed a strong belief that the statistics systems were the critical factor in maintaining volunteer interest in the project. Here, it could be seen that both before the launch of the project, and after extensive experience with running the project, the scientists and software engineers have cast the volunteers in terms of the deficit model, albeit in different forms. It was be seen that this shift in the deficit model can be seen as resulting from the project's scientists' and software engineers' attempts to resolve a number of factors, including the pursuit of their main goals (those defined by the terms under which the project has been funded, and others), accommodating and accounting for the observed behaviour of volunteers during the course of the project, fitting in with accounts of volunteer behaviour and
motivation provided by other key figures involved in BOINC-based projects, all combined with a desire to maintain authority within the project and a clear boundary between themselves and the volunteers.

Thus, although the statistics systems themselves have been largely unchanged since the project launch, their social reality and significance for the scientists and software engineers involved have changed.

climateprediction.net and Galaxy Zoo compared Chapter 6 concluded by comparing and contrasting these statistics systems with the methods by which the contributions of Galaxy Zoo volunteers are publicly acknowledged. It can be seen that differences in the approaches of the two projects have arisen from a number of factors. One is the core scientific tasks that the volunteers in the respective projects are asked to undertake. Another factor is that, as the project unfolded, Galaxy Zoo core team members started to perceive that volunteers could make important scientific contributions through serendipitous discoveries and spinoff projects; by contrast, the climateprediction.net team have not felt that that project's volunteers might be able to make contributions beyond the core task of running climate models on their personal computers, and thus have not had to reflect on the possible negative impact of statistics systems on other possible forms of volunteer contribution. A final factor was revealed by focussing on how members of the respective project teams have pursued interests outside of the confines of their own disciplines: some Galaxy Zoo core team members are seeking to pursue interests through the Citizen Science Alliance whilst climateprediction.net's team members key interest outside of climate science was demonstrating a distributed approach to ensemble modelling.

8.1.3 Chapter 7: Findings regarding climateprediction.net volunteers

In Chapter 7, I focussed on the volunteer community of climateprediction.net (thus, this can be understood as an attempt to answer RQ 3). I set out to show that the credit and other statistics systems have played a critical role in the project as experienced by a variety of volunteers. These statistics systems operate in a uniform fashion across the project and the policies that govern how they operate have been largely unchanged over the lifetime of the project. The core of my argument was that although, on the one hand, the mechanisms by which the statistics systems operate can thus be understood as largely invariant at all points of the project both spatially and temporally, their social meanings are nevertheless highly sensitive to context, in the sense that they have come to have disparate meanings and been used for different purposes across the contexts of different individuals' lived experiences of the project.

During the course of my study of climateprediction.net, a number of different classes of volunteers have been identified, based on the extent and nature of their participation in the VCPs: these different classes contribute to VCPs in different ways and therefore contribute in different ways (and to a different extent) to the pursuit of the project's various main goals and objectives. In Chapter 7, I compared and contrasted two of these classes (which I named Super-Crunchers and Alpha-Testers), explaining how both play a very important role for the project in terms of the pursuit of its goals. It was clear that both groups have sought to attain particular statuses within the climateprediction.net volunteer community. In particular, I argued that the statistics systems play very important, but very different, roles in each group's pursuit of these statuses: i.e. just as with the scientists and software engineers in both Galaxy Zoo and climateprediction.net, their interests in pursuing these statuses can be seen as socio-technically located, in the sense that it is through the affordances of the statistics systems and the way they operate that these particular volunteers have perceived the opportunity to pursue interests and have thus attempted to do so.

8.2 Reflections on public engagement with science

Here, I relate my findings to the existing literature on public engagement with science that was reviewed in Chapter 2. In subsection 8.2.2 below, I consider how my findings relate to interactions amongst laypeople around issues of technoscience. However, first I turn to how they address forms of engagement between the technoscientists and lay volunteers.

8.2.1 Interactions between laypeople and scientists: Challenging the deficit model?

In Chapter 2, Section 2.2, I grouped the nature of the relationships between laypeople and technoscientists that has been discussed in the existing literature into three broad categories.

One category involves situations where it is very much in the interests of those technoscientists interacting with the public to adopt the discourse of members of the public. These include situations such as popular science publications, media reporting of science and technology, and museum exhibitions. In these situations, the scientific and technological experts display a great deal of apparent social learning, which might be understood both as resulting from a need to engage successfully with laypeople (e.g. the commercial need to sell newspapers or books) and that these experts have become enculturated in institutions with a long tradition of successful public engagement.

As was seen in this thesis, the scientists and software engineers involved CCP are usually not embedded in institutions in which they will have learnt cultural habits (or received extensive training) regarding the successful communication of scientific knowledge to the public, and this is a critical difference between the contexts of CCPs and the situations in this category.

The other two categories are those in which scientists have had little or no prior experience of interacting with the public. The first of these two categories involve situations where technoscientists attempt to construct and maintain boundaries between themselves and members of the public In these discursive moves, technoscientists have been found to use a range of rhetorical (e.g. Wynne 1982) and material (e.g. Kaminstein 1996) tools to cast lay members of the public in terms of what has been labelled the ‘deficit model’ (Miller 2001). In this model, members of the public who are not certified as professional scientists or technologists are cast as lacking the relevant expertise (regarding scientific knowledge, modes of reasoning) and disinterestedness to evaluate issues of technoscience or contribute to the making of technoscientific knowledge.

Wynne (and others) have demonstrated how some members of the public can possess forms of expertise that are relevant (and potentially very useful) to these processes, but casting these laypeople in terms of the deficit model serves to cast their particular forms of expertise as invalid (Wynne 1982). This results in the exclusion of members of the public from making substantive contributions to technoscientific knowledge- and decision-making processes. It is suggested that this resistance to opening up the scientific process to lay people is motivated to a large extent by a desire on the part of technoscientists to conceal the messiness and indeterminacy of this process and the limits of scientific knowledge.

The other category identified in Chapter 2 involve instances where some lay members of the public or lay groups have been fully-integrated into, and make substantive impacts upon, at least some stages of technoscientific decision-making processes (for instance, Epstein 1996; Kerr *et al*. 2007). Here, it appears that some technoscientists have allowed the lay-science boundary to be breached, but in reality, the members of the public allowed to breach this boundary have adopted the discourses of science, and hence technoscientists have been able to continue to construct a boundary, conferring legitimacy to contribute on those within the boundary (i.e. those who deploy scientific discourse) and denying this legitimacy to those outside the boundary.

In short, then, studies presented in the social science literature on public engagement of science have tended to technoscientists as either:

1) Excluding members of the public from technoscientific decision-making and knowledge-generating processes by engaging in discursive moves to construct lay people in terms of the deficit model; or
2) Opening-up these processes to certain members of the public, who are able to contribute their expertise, but only once these laypeople have developed a high degree of *'cultural competence'* (Shapin 1990, p. 993) in scientific discourse related to the issue under consideration.

The work presented in this thesis, however, suggests forms of public engagement in science that differ from the forms considered in Chapter 2, with lay volunteers in CCPs able to contribute to (or have influence over) technoscientific knowledge- and decision-making processes without necessarily achieving a particularly high level of cultural competence in the relevant technoscientific domains. In this Section, I will consider how, in my case studies of CCPs, some technoscientific processes have been opened up to lay volunteers, some of the ways in which the recognition of particular forms of expertise possessed by volunteers have been recognised by technoscientists and the conditions under which they have been allowed to contribute this expertise.

Studies of Expertise and Experience (SEE): In order to gain a better understanding of this, I will draw upon some of the concepts introduced as part of the Studies of Expertise and Experience (SEE) programme, which followed the publication of Collins & Evans' paper 'The Third Wave of science studies: Studies of experience and expertise'

(2002). Informed by the work of scholars of public engagement of science such as Bryan Wynne, Collins & Evans argue that the democratic legitimacy of such processes requires the substantive involvement of non-technoscientists. However, they also argue that these processes would be undermined by simply dissolving the boundaries between scientists and non-scientists; instead, different stakeholders will possess different forms of expertise and experience which will have a bearing upon the nature and extent of what it is appropriate or desirable for them to contribute to a particular technoscientific decision-making process.

The resultant aim of Collins & Evans has been to develop a normative theory of expertise to help inform the design and conduct of such processes. This has given rise to a number of case studies to identify different forms of expertise, particularly focussing on interdisciplinary scientific collaboration (for instance, Schilhab (2007) and Weinel (2007)), and consideration of how they might interact to produce technoscientific knowledge (Collins 2007).

Here, I will highlight two particular classes of expertise that have been identified (Collins & Evans 2007), as they will be relevant in the discussion below:

1) ***Specialist tacit knowledge.*** This refers to expertise that goes beyond knowledge of what is considered scientific fact (e.g. the sort of knowledge that might be required for quizzes). Collins & Evans divide it into two subclasses:
    a. ***Interactional expertise.*** This refers to mastery of the language of a specialist domain, but without the practical competence to perform activities (e.g. laboratory work) within the domain; and
    b. ***Contributory expertise.*** This refers to the possession of practical competence within a domain.
2) ***Transmuted expertise.*** This refers not to technical or scientific expertise, but to social expertise. Possession of this form of expertise does not involve possession of specialist expertise, but rather involves the ability to make judgements about the claims of technoscientific experts on criteria such as their demeanour, the consistency of their remarks, and their social locations (e.g. institutional affiliation). This form of expertise is broken down into two subclasses:
    a. ***Ubiquitous discrimination.*** This type of discrimination is based upon general expertise gained by an individual through their experiences in society in terms of choosing between politicians, experts, salespeople etc.; and

b. ***Local discrimination.*** This involves the individual making judgments based on local knowledge about those around them.

climateprediction.net: A deficit model regarding climate science, but room for volunteers to contribute expertise about technological aspects of the project In the case of climateprediction.net, we saw that there was primarily merely a shift from one form of the deficit model to another form in terms of how the project scientists and software engineers cast the volunteers in respect of their understandings of climate science. Wynne (2006) suggests that the deficit model is very durable in encounters between technoscientists and lay people, to the extent that the former are liable to engage in very deft moves to maintain the deficit model (and thus their pre-eminent position in technoscientific decision-making processes) even in the face of pressure to open up these processes to substantive contributions from lay people, and this seems to have been the case with climateprediction.net.

Nevertheless, it should be noted that the technoscientists did recognise some forms of expertise on the part of some volunteers, albeit in very limited domains primarily relating to the technological infrastructure of the project (rather than the domain science). For instance, Alpha-Testers are recognised as possessing contributory expertise with regards to the development of the project's technological infrastructure. Another example relates to when a volunteer encounters difficulties in installing or running the BOINC middleware or climateprediction.net models on their computer: rather than providing a service to deal with volunteers' enquiries themselves, the project's software engineers instead encourage volunteers to seek the advice of other volunteers on the project's forums. Implicit in this is a recognition that some volunteers possess the relevant contributory expertise to understand the technical workings of the project (e.g. what particular error messages might mean) and the workings of personal computers to a sufficiently advanced point where they are usually able to resolve a volunteer's query.

In both cases, it can be seen that the recognition of these expertises have been driven to a large extent by pragmatic considerations on the part of the project's technoscientists. The Alpha-Testers provide a very willing, inexpensive pool of volunteers with the wherewithal to test the project's software, identify bugs, and so on. In the case of the forums, encouraging volunteers to ask other volunteers for assistance was seen as a solution to being able to deal with the sheer number of technical queries that volunteers generate: prior to the launch of the forums, there was a contact form for volunteers to ask the project's software engineers for help, but the software engineers soon became overwhelmed by the volume of queries they received and thus volunteers with problems are now directed to ask other volunteers for assistance forums.

Galaxy Zoo and volunteers' contributory expertise to the project's core work Unlike climateprediction.net, participating in Galaxy Zoo does not require the installation and running of middleware on a volunteer's computer, and thus there is reduced scope for volunteers to contribute expertise on the forums in respect of such technological aspects of the project. However, Galaxy Zoo has instead opened up some scientific processes to contributions from volunteers, as is reflected in the Galaxy Zoo ethos and the way in which it casts into being a gift economy of knowledge involving both project scientists and volunteers. In particular, volunteers have been able to contribute galaxy classifications (to contribute to the project's core work) and to identify possible serendipitous discoveries.

In terms of the project's core work, the volunteers work according to a schema that is very rigidly-defined by the project scientists. Whilst volunteers are recognised by the project scientists as possessing contributory expertise in respect to simple classificatory tasks, they are not recognised as possessing expertise in respect to other aspects of the core work (e.g. in terms of formulating research questions, interpreting results, and so on): thus, the recognition of their contributory expertise is restricted to a narrow domain. It was seen in Chapter 4 that the opening-up of the process of galaxy classification to laypeople was driven by disciplinary interests (improving the status of cosmology in relation to other subdisciplines of astronomy through the production of large-scale statistics about galaxy morphology), and in the absence of other feasible methods for producing these statistics. In other words, with respect to the core work, the scientific process appears to have been opened-up to volunteers' contributory expertise to the extent that the scientists felt was necessary to advance disciplinary interests, but no further.

Galaxy Zoo and serendipitous discoveries: Who contributes to framing processes? It might also be argued that, even if lay volunteers are allowed to make certain limited contributions to Galaxy Zoo, the scientists retain what Wynne (2003) calls 'propositional hegemony' (p. 401). In other words, the scientists retain complete authority over the framing of the project's core scientific work. This framing will include: decisions about which research questions to pursue; how the research is to be executed, including how the volunteers' tasks are defined; what the research outputs are; and how the research outputs are to be used.

However, in this thesis, it has been seen that scientists do not necessarily retain full authority over framing processes in respect to the other major aspect of Galaxy Zoo's work, namely serendipitous discoveries. For instance, the Galaxy Zoo core team have given a certain degree of autonomy to volunteers who are involved in many of the spin-off projects resulting from these, granting them opportunities to shape research questions and deciding how best to proceed in answering them[50]. This reflects some recognition of contributory expertise of lay volunteers by project scientists in some of the framing processes of science.

It was seen in Chapters 4 and 5 that making space for this aspect of volunteers' contributory expertise has been seen by project scientists as contributing to the success of the project (e.g. in terms of enhancing its credibility in the eyes of other astronomers, as a recruiting tool for volunteers) and to the future success of the Citizen Science Alliance (e.g. serendipitous discoveries providing a reason for continuing the project in the face of technological developments): i.e. this is another example where opening up to lay contributions appears to have been undertaken for pragmatic purposes by the project scientists.

Of course, the project scientists still retain a great deal of authority in these cases, as they are able to choose which, of the potential paths of research suggested by volunteers to give their support and attention to. This authority rests in large part because the Galaxy Zoo scientists are perceived by volunteers and scientists (both within and outside the project) alike as possessing expertise lacked by volunteers, including: understanding which potential paths match well with the priorities of the broader discipline; knowledge about the analytical tools required advance research in a way that is meaningful; and presenting results in a way that are accepted by the broader discipline; and how to secure resources to investigate the phenomenon further.

Galaxy Zoo and cosmology: Recognition of volunteers' contributory expertise by the broader academic community As discussed in Chapter 4, subsection 4.3.1, the contributions of Galaxy Zoo volunteers have formed the basis of a number of research outputs (articles in astronomical journals, conference papers, dataset releases) intended to be read, used and cited by cosmologists from outside the project. This eventual acceptance and use of these outputs strongly suggests a recognition by the broader astronomical community, i.e. beyond the project's scientists, of the contributory expertise of the volunteers in respect of the particular tasks involved.

As discussed in Section 4.2, astronomy is a special case in terms of lay involvement, as it has a rich history of amateurs making serendipitous discoveries: thus, the broader acceptance of, and interest in, discoveries such as Hanny's Voorwerp or the green peas is perhaps unsurprising.

Ultimately, too, however, many in the wider astronomical community accepted the credibility of the statistical results based on volunteer classifications. In other words, the wider community recognised contributory

[50] There was not space in this thesis to consider these, and other, issues related serendipitous discoveries in Galaxy Zoo: this suggests that studying serendipitous discoveries further might be a promising path for future research (see subsection 8.4.2 below).

expertise on the part of volunteers in relation to the production of these results. However, as was seen in subsection 4.4.2, there was initially a certain amount of resistance to accepting Galaxy Zoo volunteer classifications as credible, and project members had to engage in a great deal of work to secure this credibility. This helps to highlight that the acceptance of lay contributory expertise by the broader astronomical community is limited to a very narrow domain, namely to simple tasks that are rigidly defined by professional astronomers. Furthermore, the Galaxy Zoo astronomers have had to act as *translators* between the volunteers and the broader cosmological community, engaging in a great deal of work (including data processing and bias-testing) to transform the classifications into outputs that are meaningful and credible to other cosmologists.

Thus, it can be seen that, in terms of the core work of Galaxy Zoo, there has been some acceptance of the opening-up of the scientific process to lay people by cosmologists from outside of the project, albeit to a very limited extent. In particular, volunteers have not had to demonstrate full cultural competence in cosmology in order for their contributions to be accepted as scientifically-valid by cosmologists, but this acceptance has only been realised once translation has been undertaken by a group (the Galaxy Zoo astronomers) who are already regarded by other astronomers as possessing this competence.

Galaxy Zoo, climateprediction.net and volunteers' transmuted expertise: Determining which CCPs are successful Another way in which professional technoscientists might be thought of as having ceded authority to lay volunteers stems from the fact that, for a CCP to be successful in producing scientific output, its technoscientists must be able to attract and retain volunteers. In other words, in addition to the actors who determine the success of more traditional scientific projects (e.g. funding bodies, technologists and scientists who might be enrolled into working in the project or collaborating with the project, and so on), volunteers also have a critical say in whether a CCP is successful. Although this is also a challenge that has faced off-line projects as well, the abundance of CCPs online and the ease with which a volunteer is able to switch participation from one project to another (e.g. the BOINC interface allows participants to reallocate their computers' resources at the click of a mouse, whilst a single Zooniverse account allows a volunteer to participate in a range of projects) means that there is a real competition between projects to attract volunteers.

Effectively, this process involves the volunteers being required to use their *transmuted expertise* to choose which particular CCPs to participate in (and whether to continue participating in a project), thus promoting these projects' success. A volunteer might make their choices on the basis of *ubiquitous discrimination*, making judgments based on their understandings or beliefs that have developed as a result of their general experiences and observations of technoscience during the course of their lives. These beliefs might relate to (amongst other things): the credibility of the institutions involved in a particular project; the relative importance of the scientific discipline and trustworthiness of scientists in that discipline; whether the work of the project will contribute to important work within the discipline; and the demeanour and behaviour of scientists involved in a particular CCP (e.g. is this in accordance with the volunteer's expectations of how a scientist should conduct themselves?).

A volunteer may also make their judgments based on *local discrimination*. This might include knowledge gained from experiences in a particular CCP as they get to know and trust (or otherwise) the technoscientists involved in that project (e.g. through blogposts, interactions on the forums, or even in person) or whether they find participating in that particular CCP to be enjoyable.

It should be noted, however, that fact that the ultimate success of a CCP in terms of recruiting volunteers to perform scientific work is so dependent upon the transmuted expertise of volunteers does not result from a normative judgment by any one individual or institution that these volunteers are best placed to judge which CCPs are more deserving of support than others. Instead, the fact that these judgements have been opened up to volunteers is borne out of necessity: in a democratic society, volunteers cannot be coerced into choosing one project over another; instead, the ultimate decision about which CCP(s) to get involved with is necessarily the choice of the volunteer themselves.

Thus, again, the opening up of processes to the expertise by possessed by volunteers can be seen as a pragmatic move: allowing potential volunteers to play a critical role in deciding which CCPs will succeed is the necessary price to be paid by professional technoscientists if they wish to make use of the resources of members of the public to perform scientific work.

Opening up the 'black-box' of the scientific process It has been seen in this thesis that this need to attract and retain volunteers has, in turn, driven the opening-up of the black box of the scientific process, at least to a limited extent. For example, the Galaxy Zoo blog was introduced in a large part because the project scientists believed it would help to improve volunteer retention rates. Through this blog, the scientists involved have sought to make visible to volunteers the processes by which galaxy classifications are transformed into scientific results, the rewards and frustrations of getting these results disseminated (e.g. published in journals), and to inform volunteers of interim results.

However, the process of opening-up does not necessarily lead to complete transparency. Indeed, as my study of Galaxy Zoo has indicated, even when technoscientists appear to be opening-up their work to laypeople, there is nevertheless a great deal that may remain invisible. Technoscientists may be reluctant to open up more than they perceive to be necessary for their own purposes (for instance, one reason the Galaxy Zoo scientists do not blog about unpromising lines of inquiry is to avoid giving the impression to volunteers that their contributions to the project are in vain, which might jeopardize the retention of these volunteers). Nevertheless, the degree of openness of Galaxy Zoo scientists does raise the possibility that the messiness of the scientific process, the indeterminacy of science knowledge, and the identity of the scientists as social actors with interests might be exposed. It is issues such as these that Wynne (and others) argue that scientists' castings of laypeople in terms of the deficit model have been designed to conceal: thus, in this respect, it can be seen that the scientist-lay interaction in Galaxy Zoo deviates from the deficit model.

In summary, then, it can be seen that there has been an opening-up of technoscientific processes to laypeople in my case studies, both in terms of making space for them to contribute or use their expertise and also making more visible some of the processes undertaken by scientists. This opening-up has occurred even in cases where volunteers have not demonstrated a high level of cultural competence in the relevant scientific domain. In my case studies, contributory expertise on the part of volunteers has been recognised by project technoscientists in respect both to technological aspects of the project (in the case of climateprediction.net) and simple scientific tasks (e.g. classifying galaxies or recognising potential anomalies in galaxy images). Indeed, this expertise has also received increasing recognition by scientists from outside of the projects, for instance astronomers who cite journal articles that are based on the Galaxy Zoo datasets. Volunteers have also been called upon to use what transmuted expertise they possess to make judgments about which projects they will join. This, in turn, has helped to drive processes of opening-up the black box of scientific work, for instance to encourage volunteer retention.

However, it can also be seen that the boundaries between CCPs' technoscientists and volunteers are not completely dissolved: indeed, the opening-up appears to be limited to the extent that CCP technoscientists perceive as necessary or likely to contribute to their goals (either in terms of their project's success or other goals, for instance the success of the Citizen Science Alliance).

8.2.2 Interactions amongst laypeople

Another key finding of this thesis relevant to the literature discussed in Chapter 2 is the heterogeneity of the volunteers in CCPs. As I discussed, there have apparently been few, if any, studies of the dynamics amongst lay people around issues of technoscientific decision-making, e.g. in public engagement mechanisms. Understanding such dynamics is crucial for the construction of robust mechanisms, in order to ensure that all lay participants' voices are heard and that none are marginalised unjustly by other lay participants.

What was found in Chapter 7 was that different volunteers may enter a CCP with very different expectations of what they might get out of participation (or develop these expectations during the course of this participation). In particular, in the two groups that I focussed upon in climateprediction.net, it was found that they were very interested in pursuing social status within the community of volunteers, differentiating themselves from other volunteers.

Alpha-Testers In the case of the Alpha-Testers, it was found that they generally had educational and/or professional backgrounds in areas of technology and science relevant to climateprediction.net, and that they saw themselves as having the legitimacy to speak on behalf of the project scientists and software engineers whilst attempting to suggest that other volunteers lacked such legitimacy. In the particular instance on the project forum that was studied in Chapter 7, subsection 7.2.3, this led to conflict with other volunteers. In the case of public engagement mechanisms in general, my findings here suggest that there may be participants who, whilst being regarded as being lay (i.e. they are not part of the group of identified professional scientists and technologists), nevertheless possess professional and/or educational backgrounds relevant to the areas under discussion and feel that this confers a greater authority on themselves compared with other lay participants. What this suggests, then, is that whilst Wynne (2006) and others discussed in Chapter 2 have focussed on how technoscientists may attempt to silence or delegitimise (and thus alienate) lay members of the public, there is also the risk that the actions or discourses of some lay participants may act to alienate other lay participants in public engagement mechanisms.

Super-Crunchers The other group studied in Chapter 7 were the *Super-Crunchers*, who pursued greater social status through the acquisition of credit scores and league table rankings (either personally or through a team). Although those I studied were keen to help science through their participation in CCPs, the choice of which particular project (and thus which area of science) they were helping at any one time was usually primarily determined by the quest for improved scores and rankings rather than seeking to help one particular project or discipline.

Statistics systems and the pursuit of status Finally, it was seen that the credit and statistics system of climateprediction.net have been used by both groups in very different ways in order to pursue their respective goals regarding social status. In the case of the Super-Crunchers, the climateprediction.net scientists and software engineers had not initially anticipated that the credit and statistics systems would be used by (or indeed, of any importance to) this group of volunteers. In the case of the Alpha-Testers, these systems have been used in ways of which the project team members have not at all expressed any awareness. Relating this to public engagement mechanisms, this suggests that features and policies of these mechanisms, although possibly appearing to be spatially and temporally, may nevertheless serve a multiplicity of purposes across the various social worlds of the different participants. In particular, these features and policies may be used by lay participants to situate themselves in relation to others (or to silence or delegitimise others etc.) in ways that were unanticipated or may not be visible to others.

8.3 Reflections on Citizen Cyberscience

As explained in Chapter 1, subsection 1.1.1, projects whose primary scientific goal is the processing of data can be grouped into two broad categories, namely low-granularity projects, in which volunteers play a somewhat passive role with data being processed in the background on their personal computers (e.g. BOINC-based projects), and high-granularity projects, in which volunteers are asked to perform active tasks of data processing, such as galaxy classifications.

In subsections 8.3.1 and 8.3.2, I will start by considering both high- and low-granularity projects. In subsection 8.3.1, I will focus more on issues to which scientists interested in running a Citizen Cyberscience project might be sensitive, rather than providing concrete recommendations. In subsection 8.3.2, I will consider the Alpha-Testers (encountered in Chapter 7, subsection 7.3.2), explaining why they might be valuable to both high- and low-granularity projects, and presenting some recommendations for practice to this end. Then, in subsection 8.3.3, I will focus on low-granularity projects only, and explain how Super-Crunchers (encountered in Chapter 7, subsection 7.3.1) are particularly valuable to this type of project. From my empirical research, I draw out some recommendations for how a CCP might attract and retain such volunteers.

8.3.1 General issues in Citizen Cyberscience

As discussed in Chapter 1, subsection 1.1.2, the social scientific literature produced to date that focusses on Citizen Cyberscience problematizes neither the role of a project's team members (scientists and software engineers) nor of a project's infrastructure (e.g. features included on its website, policies for acknowledging volunteer contributions etc.). What has been seen in this thesis, however, is that these individuals are every bit as much social actors as the volunteers. They are embedded in institutional, historical, and social contexts, and have interests in these contexts (such as career advancement), and that these interests are indeterminate in that they are subject to constant negotiation and possible change over time. Furthermore, we have seen how these interests are socio-technically located, and thus can shape (and in turn shape) a CCP.

Scientists' interests and project infrastructure The team members involved in a CCP will make a constant series of decisions regarding how to set up, run, and further develop their project. They should be aware that these decisions can have a profound impact upon the relationships they build with volunteers, with the broader scientific community, and with society-at-large, and, in turn, on the pursuit of their own interests. Some decisions can help these team members to pursue their existing interests more easily, close down or render more difficult the pursuit of others, or open up new possibilities. Thus, it is very important for scientists and software engineers embarking upon (or currently involved in running) a CCP to work out in advance, and frequently reflect upon as the project unfolds, precisely what it is they hope to achieve in order that they can build their project's infrastructure to help them better to attain their goals. For instance, what particular scientific goals do they wish to achieve? To what extent do they wish to engage members of the public (e.g. the scale of engagement, the level of engagement etc.)? What other goals might they have, and in which domains are these (for instance, are they interested in success in their own scientific discipline only? Like many members of the Galaxy Zoo core team, are they keen to make a name for themselves in Citizen Cyberscience at-large? Or, like key members of the climateprediction.net team, are they seeking to establish priority for themselves as originators of novel techniques, e.g. for running mathematical simulations?).

Issues in feeding back to the volunteers Of critical importance to the success of a CCP, in particular, is how it feeds progress back to the volunteers. These volunteers give up time and effort (as well as material resources) to participate in CCPs, and can therefore be understood as being eager to ensure that these contributions have not been in vain. Furthermore, it can only be right and courteous to inform volunteers that this is the case.

Those involved in setting up and running such projects have a virtually limitless range of options regarding the features and policies they might implement for doing so. Even those working with a pre-existing structured framework, such as BOINC or projects developed in the Zooniverse, have an enormous flexibility. For instance, in the case of BOINC-based projects, policies regarding the attribution of credit can vary across projects (for instance, whether credit should be awarded only upon completion of a whole work unit or at interim stages, or whether credits are updated automatically or awarded on an ad-hoc basis, or how a project may award credit for work done on different sub-projects). In Zooniverse projects, there are also many ways in which volunteers contributions might be acknowledged, for instance whether particular volunteers be singled-out. In both cases (and other CCPs), the scientists and software engineers involved can exercise a great deal of discretion regarding how much of their own work and practices are opened up to the volunteers (e.g. presenting interim results, decisions being made regarding future developments of the project etc.) – for instance, all Zooniverse projects have a project blog, all BOINC-based projects have a news page, whilst other projects may be able to embed a blog on their website or link to another blog or Twitter feed.

Whether to implement systems of volunteer statistics In this thesis, we saw two projects that have ended up with very different approaches to the question of whether to award credit (and other) scores to volunteers. In the case of Galaxy Zoo, a high-granularity project, the core team members expressed concern that a league table of the number of galaxies classified would compromise the quality of classifications; by contrast, the speed at which a volunteer's model in climateprediction.net might be processed is not supposed to impact upon its end result, and thus the existence of a credit system is not regarded as compromising the quality of results.

It might therefore be seen that there should be a direct correspondence between the level of granularity of a project, and whether volunteers' contributions receive some form of numerical score. However, this is not so straightforward. In the case of high-granularity projects, it may be the case that it might be possible to score the quality of a volunteer's contribution. For instance, in the case of the FoldIt project (http://fold.it (accessed 6 January 2012)), the key scientific goal is to find the lowest-energy conformations of various proteins. To this end, volunteers are presented with a protein, and challenged to fold it into the lowest energy conformation that they can, and are scored accordingly.

Similarly, it is not the case scoring volunteers is necessarily the most advisable path to be taken in a low-granularity project. For instance, a project may be concerned to engage as many volunteers for as long as possible (e.g. for the purposes of educating them about scientific issues) and there may be concerns that a statistics system might send out a message that only volunteers able to process relatively large quantities of data are truly valuable to the project, and discourage other volunteer from continuing to participate. On the other hand, a statistics system might prove highly-motivating to some volunteers (e.g. Alpha-Testers) and thus might be included if a project is particularly interested in processing a large quantity of data – this is dealt with in more depth in subsection 8.3.3 below.

Whether and when to blog Another way of feeding back to volunteers is through a project blog. This may be particularly attractive to a project's scientists and software engineers if they have decided not to implement a score system (and thus might feel the need to offer other forms of feedback to volunteers), or if they anticipate that processing the volunteers' data and having it published in journal articles is a long process (and thus the volunteers might need reassuring at regular intervals that their contributions are indeed being used in genuine scientific research).

However, the decision to blog is not so straightforward. Choices must be made regarding the topics to blog about, how frequently to blog, and the content of the blog. For instance, how regular should blogposts be made in order to reassure volunteers? How open should the scientists be about their work (a key issue here relates to the project's overall goals, e.g. a project wishing to educate me the public about scientific issues and processes might see greater openness as contributing to this goal, whilst another project may be more open in order to reassure volunteers by demonstrating how their contributions are being used). However, scientists should also be aware that being very open can carry its own risks. For instance, in the case of Galaxy Zoo, the project was scooped to the first publication of results about Hanny's Voorwerp by another team of astronomers who had been originally alerted to the Voorwerp by a post on the Galaxy Zoo blog. As the discovery of Hanny's Voorwerp had already attracted substantial interest, both in the popular media and in the field of astronomy, the project has nevertheless been able to secure its priority for this discovery. This might not be the case for projects with a lower profile, and scientists and software engineers should bear this, and other issues, in mind when making decisions about blogging practices.

8.3.2 Alpha-Testers and both low- and high-granularity projects

In this subsection, I consider one of the groups studied in Chapter 7, namely the Alpha-Testers, explaining briefly why they might make important contributions to the success of CCPs, and some recommendations for practice for recruiting and retaining them. These volunteers appear to enjoy the status they have within a project, in their case, being able to align themselves with those running the project rather than the other volunteers. The way in which the credit system operates in climateprediction.net apparently provides them with a device to engage in moves to distance themselves from the volunteers, by suggesting that they themselves are unconcerned by the acquisition of credits and that this contrasts with the other volunteers, who they suggest are motivated primarily by the acquisition of credits.

The value of Alpha-Testers Although I only identified and studied these volunteers in the context of a low-granularity project, Alpha-Testers can prove valuable to high- and low-granularity projects alike, for instance in terms of providing a steady initial base of volunteers when a project is first launched or for providing feedback regarding proposed new features or developments of a project. Thus, CCPs are likely to benefit both from recruiting such volunteers from the early days of the project and retaining their interest as the project unfolds.

Recommendations for retaining Alpha-Testers If a CCP is to remain of interest to Alpha-Testers, then project team members may consider providing ways which give Alpha-Testers the opportunity to differentiate themselves from other volunteers, but in ways that do not alienate other volunteers. These might include:

a) **Publicly acknowledging the contributions of Alpha-Testers, for instance by listing them on a project's website or naming them in other public ways as in scientific articles.** Care, however, should be taken here by considering whether such an approach might be suitable for an individual project. For instance, as we have seen in Galaxy Zoo, the core team perceive the importance of treating all volunteers as though they have made equally-valuable contributions to the core work of the project and thus it may not be appropriate to single out Alpha-Testers for additional acknowledgement;

b) **Creating a badge which can be attached to Alpha-Testers' posts on the project's forums and which indicates their contribution to alpha- and beta-testing.** Alpha-Testers are frequently very active on the forums (presumably because these are the main arenas for them to be able to pursue social status): thus, the forums seem a promising location to make the Alpha-Testers' contributions more visible to other volunteers;

c) **Assigning Alpha-Testers other special responsibilities in the project**¸ for which they can receive acclaim and recognition. For instance, the Galaxy Zoo project has a volunteer-maintained online library of all publications stemming from the project: a project may choose to assign responsibility to some Alpha-Testers to perform a similar role.

8.3.3 Super-Crunchers and low-granularity projects

As discussed above, there is some concern that encouraging volunteers to complete tasks as quickly as possible in high-granularity projects might compromise the quality of a project's scientific output. However, in the case of low-granularity projects, such as BOINC-based projects, this is not really a concern and such a project might be particularly advised to ensure that it recruits a corpus of Super-Crunchers. To recap, this class of volunteers is characterized by the relatively large quantity of project data that they process on their own computers. As a result of this, the Super-Crunchers participating in such a project can be seen to make a particularly valuable contribution to the project's goal of producing publishable scientific results: (for instance, approximately 60 per cent of all credit awarded by climateprediction.net has been awarded to the Super-Crunchers).

The motivations of Super-Crunchers In Chapter 7, subsection 7.2.2, I focussed in particular on Super-Crunchers who are involved in teams. Holohan & Garg (2005) argued that such volunteers (who I have named Team Crunchers) are primarily motivated by what they termed *'co-opetition'*, in which they are keen to promote their team's credit scores and improve their team's league table positions across a range of BOINC-based projects in competition with other teams, but co-operate within their teams in order to help other team members improve their own credit scores (and thus improve the team's scores and league table positions). I extended this analysis, by arguing that there is a form of *nested co-opetition*, in that Team Crunchers are indeed motivated by inter-team competition at the level of individual BOINC-based projects, but that they are also motivated by competition within their own team. This competition may take the form of league tables of individual team members, or involve the formation of sub-teams who compete with each other (and whose members co-operate with each other). I will now attempt to distil some recommendations for practice from these findings.

Recommendations for retaining Super-Crunchers climateprediction.net appears to have been successful in retaining a corpus of Team Crunchers, and it is very likely that this is due in large part to the way that the project's credit system and online forums operate. Its credit system has been relatively stable and consistent, and furthermore is updated on a daily basis, so that volunteers have a high degree of trust in its reliability and robustness as an accurate measure of work completed. This means a number of things. One is that when a team reaches a particular credit milestone or league table positions, other volunteers are willing to offer praise and congratulations because they are confident that it represents a genuine accomplishment. Another is that, should a team turn its attention to climateprediction.net and particularly encourage its members to donate computation resources to the project, then this will indeed result in a swift accumulation of credits and a rapidly-improving league table position (for instance, the Dutch Power Cows team studied in Chapter 7, subsection 7.2.2, engages in so-called 'stampedes', where one project is selected and all members are asked to focus their attention on this project). Equally important here is the activity of volunteer online forums: because there are so many volunteers reading and posting on them, this reassures the Team Crunchers that there are a large number of volunteers who are interested in hearing about their team's successes. In general, if low-granularity projects wish to attract and retain the interest of a body of *Super-Crunchers* to process data, the following recommendations are made:

a) **A project should include a system of attribution/scoring (such as the credit system), along with a team system,** that is easy to understand by volunteers and regarded as an accurate and reliable indicator of a particular volunteer's contribution to the project. Such a system should:
   - Be cumulative, so that an increase in score is associated with additional data processed;
   - Be consistent throughout the duration of a project;
   - Operate under rules that are equally applicable to all volunteers, so that all volunteers are scored on the same basis, allowing for easy comparison between volunteers; and
   - Be updated on a regular basis.

b) ***A project should ensure that the scores achieved by teams are very visible to others.*** This happens within climateprediction.net owing to the very active forums and the display of league tables on its website. It is not, however, inevitable that a CCP's forums will be used by a large number of volunteers. To nurture activity, those who run a project could post regularly, for instance, about the science underlying their project, issues relating to the science to stimulate debate (the fields of many projects relate to current political debate), or make posts when teams have reached credit milestones in order to encourage a culture of praise for teams who achieve high scores.

These recommendations could follow from the work of Holohan & Garg (2005) only, as they are about promoting inter-team competition on the level of a specific project. However, my findings also suggested that competition within a team was an important motivating factor for Team Crunchers. Thus, I also make the following recommendation:

c) **A project should enable easy comparison of individuals (and possibly subteams) within a particular team.** This might be achieved in a number of ways:
   - A team leader might be given the option to export statistics for their team's members for use on their team's website; or
   - A project's website might have a page for each team registered with the project, displaying a league table only ranking that team's members. A project might also consider enabling sub-teams to register with the project and providing, for a team, a league table of its sub-

teams. Additionally, the project might allow for these league tables to be embedded on a team's website.

8.4 Suggestions for future work

One of the things that surprised me (and overwhelmed me at times) was the complexity and richness of the communities that have formed around the CCPs that I studied. As a result, what I have presented in this thesis is necessarily only a very narrow slice of what has occurred, and there is tremendous scope for further work on these projects that might prove of interest to a range of people, including scientists and software engineers involved in CCPs, and scholars in STS (both those interested in public engagement with science and those with other research interests). Here, I discuss two areas of these projects that would form a promising basis for future research, namely studying groups of volunteers other than those studied in Chapter 7, and a focus on serendipitous discoveries (see Chapter 4, subsection 4.3.2). I have already collected some data that would assist in this research, and a more concerted effort (both in terms of data collection and analysis) would bring studies of both of these areas to fruition.

8.4.1 Other groups of volunteers

As I discussed in this thesis, neither pre-existing social studies of CCPs (see Chapter 1, subsection 1.1.2) nor social studies of lay participation in technoscientific decision-making processes (see Chapter 2, Section 2.3) have studied how different groups of lay people may interact with each other around issues of technoscience. Chapter 7 presented some work that addresses these gaps in the literature. However, here I focussed only on two particular groups of volunteers (Alpha-Testers and Super-Crunchers), active in climateprediction.net.

These groups might be considered unusual, and not representative of active climateprediction.net volunteers as a whole. Together, they comprise barely 10% of all active volunteers, and both have made particular, but different contributions, to the success of the project. But what of the remaining active volunteers? In climateprediction.net, fewer than a quarter of all credits are attributed to 80% of the volunteers, and these volunteers generally appear to have little or no contact with scientists or scientific institutions other than through climateprediction.net.

Potential contribution to Citizen Cyberscience Unlike the Super-Crunchers and the Alpha-Testers, they do not appear to make a particularly important contribution to the scientific goals of the project (in terms of processing data), nor to the development of new features or iterations of the project. Nevertheless, they are very important to the project because they are the focus of its outreach efforts: climateprediction.net is a valuable way of communicating science to a group of people for whom scientific institutions may have few other opportunities to engage them. Thus, advancing understandings of how these volunteers experience a CCP, not only in terms of their interactions with the project's scientists and software engineers but also with other volunteers, is important for any project (either low- or high-granularity) whose goals include engagement with a large number of laypeople for outreach and education. For instance, in Chapter 7, subsection 7.2.3, we saw conflict between the Alpha-Testers and other volunteers: it is important for a project to ensure that tensions between volunteers are resolved or minimised in order to reduce the risk of volunteers feeling alienated by other volunteers, and leaving the project.

Potential contribution to understanding public engagement with science In addition, studying these volunteers might also prove of interest to STS scholars who focus on public engagement with science, and to those who are concerned with the design and implementation of public engagement mechanisms for issues of governance of science and technology (for instance, see Chilvers 2008; Rowe & Frewer 2004). As Wynne & Felt (2007) found, for instance, the majority of the European public have little or no contact with scientific institutions and little scientific background: if true democratization of governance processes is to occur (Ravetz 1971; Noble 1977), these mechanisms will, by necessity, include many such people. Thus, understanding how people with little prior contact with scientific institutions approach issues of technoscience, and how they interact with other people (lay and expert) around these issues, is important for the development for effective mechanisms for the inclusion of lay people in the governance of science and technology and to ensure they are not marginalised or alienated by either the technoscientists or other lay participants (for instance, see Webler & Tuler 2000).

8.4.2 Serendipitous discoveries

In my case study of Galaxy Zoo in Chapters 4 and 5, I focussed primarily on the core work of the project (galaxy classifications). However, as I explained (Chapter 4, subsection 4.3.2), serendipitous discoveries have also played an important role in the life of the project. Examining the story of Hanny's Voorwerp and/or the Green Peas in greater depth could make a contribution to a number of fields of study.

Tensions regarding the recognition of volunteers As I discussed in Chapter 5, Section 5.3, a key feature of the Galaxy Zoo ethos is that all volunteers should be treated as being of equal value and this is embedded in the methods by which volunteer contributions to the core work of the project are publicly acknowledged. However, this approach has not applied in the case of serendipitous discoveries: the first person to post an image of Hanny's Voorwerp on the project forums and ask what it is has subsequently been included as a co-author on all papers about the object, launched her own website, had a comic book written about her discovery, and, of course, even had the object named after her. In the case of the Green Peas project, a self-selecting group of ten volunteers called themselves the *'Peas Corp'* and have been listed by name in the footnotes of all papers on the topic (having been offered, and turned down, the opportunity to be listed as full co-authors).

This points to a tension which could enable the study of a number of questions of interest. How do volunteers who might have been involved in serendipitous discoveries respond to (or seek out) the prospect of being given additional recognition, and how do they negotiate and construct their own identities within and outside the project? How do other volunteers, used to an egalitarian approach to volunteer acknowledgment, perceive these moves? In a project such as Galaxy Zoo, where a gift economy of knowledge has emerged in relation to the core work of the project, does the singling out of a volunteer as having made a contribution of particular value to the project sit easily with an underlying assumption in a gift economy that all objects being exchanged are endowed with value because of the sacrifice of the donor, not its value to the recipient? Do the project scientists perceive a volunteer's self-promotion as appropriate or even desirable behaviour (e.g. in relation to Mertonian norms)? To what extent do the project scientists contest or accept the inclusion of some volunteers' names alongside their own in lists of authors of scientific articles? How is the elevation of status of some lay volunteers regarded by the broader scientific community? And many other questions besides.

Potential contribution to Citizen Cyberscience The answers to such questions might be of interest to many scientists and software engineers involved in a wide range of CCPs who are considering how to credit volunteer contributions. In many such projects (both high- and low-granularity), it would be quite straightforward to tie a breakthrough to a particular volunteer, and thus single out individuals for recognition. For instance, in the case of PlanetHunters (a high-granularity project and part of the Zooniverse – see Chapter 4, subsection 4.3.3), volunteers are asked to look at telescope images and flag up objects that might be new planets, and these are subjected to further analysis. Should an object then be certified as a planet, then it is quite straightforward to identify the volunteer who initially spotted it. In the case of ProteinFolding@Home or Rosetta@Home (both BOINC-based, and thus low-granularity, projects), the task is to test all possible folding conformations of particular proteins, in order to find the lowest-energy conformation: it is possible to trace the volunteer upon whose computer the work unit containing this conformation had been run.

In such a situation, a project's team members will need to make a number of decisions about the level of recognition to be given to an individual. On the one hand (as in the case of Hanny's Voorwerp), the idea that a lay person might be credited with a scientific discovery has proven to be a powerful tool for recruiting new volunteers. On the other, however, singling out a volunteer for praise may provoke tension or resentment amongst other volunteers and even amongst project scientists – is it fair, for instance, to credit one volunteer over something that might be considered, essentially, a matter of luck when many volunteers have also given up their time and effort to contribute? How this is resolved may play an important role in determining the eventual success (or otherwise) of a CCP.

Potential contribution to Feminist Studies of Science and Technology Another domain of scholarship to which consideration of Hanny's Voorwerp in particular might make a contribution is feminist approaches to the study of science and technology, as there are a number of parallels between the story of Hanny's Voorwerp and some strands of feminist thought. As documented in Bowler & Morus (2005), this is a diverse field. For instance, some feminist scholars have sought to understand why science (both historically, and in the present day) is a domain that is dominated by men, or have argued that women have made important contributions to science but have been marginalised in historical and social accounts (examples include Abir-Am & Outram 1987; Maddox 2002; Stein 1985). Other scholars have argued that what has been cast as the objective, scientific method in canonical accounts of science is, in fact, just one of a range of valid ways of developing knowledge about the world (Haraway 1991) and, indeed, that this scientific endeavour has, since the Scientific Revolution, involved a domination of Nature by Man (Fox Keller 1985).

These two approaches came together in Evelyn Fox Keller's biography of the plant geneticist Barbara McClintock (Fox Keller 1983), both to celebrate and bring a woman scientist to greater prominence, and to show how an approach to the study of science that, at the time, was considered unorthodox by leading members of her field, in fact made a profound contribution to the advancement of understandings of genetics.

Quite apart from the discoverer of Hanny's Voorwerp, Hanny van Arkel, being a woman, there are some interesting similarities between the discovery of the Voorwerp (and the discourse of Galaxy Zoo scientists in terms of the attitudes that they wish to cultivate amongst individual volunteers when looking at galaxy images) and the work of McClintock. At the time McClintock was working, the prevailing approach in genetics was very quantitative, with individual scientists expected to study organisms (plants, yeast, fruit flies etc.) using statistical methods, involving a large quantity of data, to identify and quantify standard genetic properties. By contrast, McClintock was interested not in the average, but in the particular, i.e. individual organisms (in her case, maize) with anomalous properties. Rather than swiftly processing large quantities of organisms in order to produce means, standard deviations etc., she would instead focus on a very small number of anomalous maize plants, being sufficiently curious to take care over each one in order to understand its particularities, i.e. quality over quantity.

Similarly, as we have seen in Chapter 5, the Galaxy Zoo core team have been keen to encourage volunteers to take time over each galaxy classification, partly to improve the quality of individual classifications but also to encourage volunteers to look out for the anomalous or unexpected and to bring it to the attention of other volunteers and the project scientists. To this end, they abolished the league table, in order to discourage volunteers from clicking through as many galaxies as possible (i.e. to promote quality over quantity) and including an instruction on the volunteer classification interface that a volunteer should post unexpected images to the project's forums[51]. In various fora (such as popular science publications, the comic book, and in the words of many project scientists in my interviews), Hanny van Arkel has been praised for being interested in the particular and anomalous, and also for her persistence in asking project scientists what the unusual object might be, i.e. her curiosity. Thus, a more detailed study of the story of Hanny's Voorwerp might prove of interest to a number of feminist scholars of science.

8.4.3 Technoscientists as lay with respect to other domains of technoscientific knowledge

In this thesis, I have focussed on problematizing both the relationships between the technoscientists involved in a CCP and the volunteers, and relationships amongst volunteers. However, apart from making a distinction between the Galaxy Zoo project's core team and other astronomers involved in the project, I did not problematize relationships amongst the technoscientists of a CCP, instead usually treating them as a homogeneous whole.

Although the technoscientists often present a quite united front to the volunteers, with each technoscientist presented as firmly on the expert side of the expert/lay divide, it must be stressed that they are usually drawn from a range of backgrounds and work in different contexts (such as disciplinary, institutional and national). In particular, CCPs are normally collaborations between software engineers and domain scientists, such as astronomers or atmospheric physicists, and these scientists may in turn be drawn from a number of different academic subdisciplines and institutions.

[51] Of course, the overall core aim of the project is to produce statistics from analysis of very large datasets, but I am focussing here on the behaviour of the individual volunteer.

Thus, it is clear that whilst a particular CCP technoscientist will be expert with respect to some of the domains of technoscientific knowledge relevant to the project, they are likely to be lay with respect to other domains. For instance, in Galaxy Zoo, a software engineer might be lay with respect to cosmology.

They may lack expertise regarding what is accepted as cosmological fact and theory, issues of debate at the cutting edge of the discipline, and the processes by which cosmologists generate knowledge is generated (e.g. how they analyse and collect data, and modes of reasoning). Conversely, a cosmologist may be unaware of the state-of-the-art in software engineering, lacking an understanding of what is technically feasible or infeasible.

However, it is also useful at this stage to bear in mind that knowledge does not simply refer to a collection of facts and theories and their production by scientists, but instead that knowledge is located in, and sustained by, socio-technical networks. Thus, to refer to somebody as lay with respect to a particular domain of knowledge, what it really means is that the individual is lay with respect to the socio-technical network in which it is located.

Studies of distributed e-Science projects (such as Olson *et al*. 2008) point to a number of aspects in which members of a CCP's team might be lay with respect to relevant domains of knowledge, in the sense Individuals can find it difficult to conceptualise contexts in which their collaborators in other disciplines or institutions are working (Hinds & Weisband 2003).

One element of these contexts is that collaborators may be embedded in different reward structures: whilst, for instance, software engineers may be seeking to build novel, ground-breaking systems in order to make their name in their discipline, domain scientists want them to build systems that best support the scientific work of the project (and thus will enable the domain scientists to advance their careers in their own discipline), irrespective of novelty (Spencer *et al.* 2008). In this example, each group may lack expertise regarding the reward structures in which the other is embedded, and thus be lay in terms of the technoscientific work that the other group is trying to accomplish within the project[52]. Other differences in context can include: the resources available to various collaborators which can constrain and shape the work they are able to conduct in the project; and existing hierarchies in each team member's home institution, in the sense that a team member may feel the need to conform to these (and conduct their work in the project accordingly) in addition to the project's lines of authority (de Rooij *et al.* 2007).

The literature here deals with e-Science projects which are not public-facing. CCPs, however, are, and this would provide a further, interesting dimension for study. Possible research questions could include: How do domain scientists interest and engage software engineers to work on a CCP? How are shared understandings

[52] In Chapters NUMBER, the relationship between the rewards that various members of the two CCPs' teams

achieved (or not) between domain scientists and software engineers regarding the project's goals and work plans? By what processes and to what extent do technoscientists develop expertise relating to technoscientific aspects of the project with respect to which they were originally lay? How do technoscientists negotiate and manage their identities across different contexts, i.e. comparisons between contexts that are visible to the volunteers (e.g. on a project's website) and those that are not (e.g. during internal decision-making processes)?

Potential contribution to Citizen Cyberscience The literature discussed this subsection provides a useful resource for technoscientists involved running CCPs, highlighting as it does a number of key challenges involved in fostering successful collaborations between software engineers and domain scientists. Indeed, in many cases, it gives rise to concrete recommendations for practice (e.g. Hinds & Weisband 2003).

The work outlined in this Section would complement this literature, as it would help to improve understandings of how to manage the issues that arise in multidisciplinary and multi-institutional scientific collaborations in projects that have a strong element of public involvement. For instance, if a project is to be truly open with volunteers about what is going on inside the project, how might they acknowledge these challenges without losing the confidence of volunteers that the project is functioning effectively?

Potential contribution to studies of technoscientific expertise The work outlined in this subsection would help to extend the analysis in subsection 8.2.1 by providing a richer picture of the processes by which various forms of expertise in climateprediction.net and Galaxy Zoo are contested, negotiated and recognised, and how these processes unfold over time. In particular, it should offer interesting insights regarding how technoscientists manage their identities as lay with regards to particular technoscientific aspects of the project, whilst at the same time possibly trying to retain authority in relation to the volunteers (for instance, how do they manage situations in which they may be non-expert, whilst some volunteers are recognised as having certain forms of expertise?).

Bibliography

Abell, J. & G. Myers (2008), 'Analyzing research interviews', in (eds.) Wodak, R. & M. Kryzanowski, *Qualitative Discourse Analysis in the Social Sciences*, New York: Palgrave Macmillan, pp. 145-160.

Abir-Am, P. & D. Outram (eds.) (1987), *Uneasy Careers and Intimate Lives: Women in Science, 1787-1979*, New Brunswick, New Jersey: Brunswick University Press.

Aiken, M. (2009), 'Wind power planning controversies and the construction of "expert" and "lay" knowledges', *Science as Culture*, 18(1), pp. 47-64.

Allen, B. (2007), 'Environmental justice, local knowledge, and after-disaster planning in New Orleans', *Technology in Society*, 29(2), pp. 153-159.

Allen, M. (1999), 'Do-it-yourself climate prediction', *Nature*, 401, p. 642.

Allen, M., M. Collins & D. Stainforth (2006), 'Method and system for producing a weather forecast', *US Patent No. 7,016,784*, Washington, DC.

Amsterdamska, O. (1990), 'Surely you are joking, Monsieur Latour!', *Science, Technology & Human Values*, 15(4), pp. 495-504.

Anderson, D. (2003), 'Public computing: Reconnecting people to science', *Proceedings of the Conference on Shared Knowledge and the Web*, Madrid, Spain, November 2003

Anderson, D. (2004), 'BOINC: A system for public-resource computing and storage', *5th IEEE/ACM International Workshop in Grid Computing*, November 8 2004, Pittsburgh, USA.

Anderson, D. & G. Fedak (2006), 'The computational and storage potential of Volunteer Computing', *Proceedings of CCGRID'06*, pp. 73-80.

Arksey, H. & P. Knight (1999), *Interviewing for Social Scientists: An Introductory Resource with Examples*, London: Sage Publications.

Bakhtin, M. (1981), 'Social heteroglossia', in (trans.) Holquist, M., (eds.) Emerson, C. & M. Holquist, *The Dialogic Imagination*, Austin, Texas: University of Texas Press, pp. 74-81.

Baldassari, J., D. Finkel & D. Toth (2006), 'SLINC: A framework for Volunteer Computing', *Proceedings of the 18th IASTED International Conference on Parallel and Distributed Computing and Systems*.

Bamford, S., R. Nichol, I. Baldry, K. Land, C. Lintott, K. Schawinski, A. Slosar, A. Szalay, D. Thomas, M. Torki, D. Andreescu, E. Edmondson, C. Miller, P. Murray, M. Raddick & J. Vandenberg (2009), 'Galaxy Zoo: The dependence of morphology and colour on environment', *Monthly Notices of the Royal Astronomical Society*, 393, pp. 1324-1348.

Barbot, J. & N. Dodier (2002), 'Multiplicity in scientific medicine: The experience of HIV-positive patients', *Science, Technology & Human Values*, 27(3), pp. 404-440.

Barnes, B. (1982), *T. S. Kuhn and Social Science*, London: MacMillan.

Barnes, B., D. Bloor & J. Henry (eds.) (1996), *Scientific Knowledge: A Sociological Analysis*, London: Athlone Press.

Barnes, B. & R. Dolby (1970), 'The scientific ethos: A deviant viewpoint', *European Journal of Sociology*, pp. 3-25.

Barnes, M. (1999), *Building a Deliberative Democracy: An Evaluation of Two Citizens' Juries*, London: IPPR.

Barton, R. (2003), ' "Men of science": Language, identity and professionalization in the mid-Victorian scientific community', *History of Science*, 41, pp. 73-119.

Bauer, M., N. Allum & S. Miller (2007), 'What can we learn from 25 years of PUS survey research? Liberating and expanding the agenda', *Public Understanding of Science*, 16(1), pp. 79-95.

Baym, N. (1995), 'From practice to culture in Usenet', in (ed.) Star, S., *The Cultures of Computing*, Oxford: Blackwell, pp. 29-52.

Beatini, M., P. Gay, B. Keel, K. McCullough, M. Schoenberg, Jason Thibeault & Jodi Thibeault (2010), *Hanny and the Mystery of the Voorwerp*, Edwardsville, Illinois: Citizen Science Alliance.

Beatty, J. (2006), 'Masking disagreement among experts', *Episteme*, 3, pp. 52-67.

Beberg, A., D. Ensign, G. Jayachandran, S. Khaliq & V. Pande (2009), 'Folding@home: Lessons from eight years of volunteer distributed computing', *2009 IEEE International Symposium on Parallel & Distributed Computing*, pp. 1-8.

Beck, U. (1992), *Risk Society: Towards a New Modernity*, London: SAGE Publications.

Benckler, Y. (2006), *The Wealth of Networks: How Social Production Transforms Markets and Freedom*, New Haven: Yale University Press.

Bennett, P. & S. Smith (2007), 'Genetics, insurance and participation: How a Citizens' Jury reached its verdict', *Social Science & Medicine*, 64, pp. 2487-2498.

Biagioli, M. (1998), 'The instability of authorship: Credit and responsibility in contemporary biomedicine', *FASEB Journal*, 12, pp. 3-16.

Biagioli, M. (2003), 'Rights or rewards? Changing frameworks of scientific authorship', in (eds.) Biagioli, M. & P. Galison (2003), *Scientific Authorship: Credit and Intellectual Property in Science*, New York: Routledge, pp. 253-279.

Bourdieu, P. (1988), *Homo Academicus*, London: Polity.

Bowler, P. & I. Morus (2005), *Making Modern Science: A Historical Survey*. Chicago: Chicago University Press.

Brenna, B. (2011), 'Clergymen abiding in the fields: The making of the naturalist observer in eighteenth-century Norwegian natural history', *Science in Context*, 24(2), pp. 143-166.

Brooks, H. (1968), *The Government of Science*, Cambridge, Mass.: MIT Press.

Bucchi, M. (1998), *Science and the Media: Alternative Routes in Scientific Communication*, London: Routledge.

Buchanan, E. (2011), 'Internet research ethics: Past, present, and future', in (eds.) Consalvo, M. & C. Ess, *The Handbook of Internet Studies*, Oxford: Wiley-Blackwell, pp. 83-108.

Burawoy, M. (ed.) (2000), *Global Ethnography: Forces, Connections and Imaginations in a Postmodern World*, Berkeley: University of California Press.

Burri, R. (2008), 'Doing distinctions', *Social Studies of Science*, 38(1), pp. 35-62.

Callon, M. (1986a), 'Some elements of a sociology of translation: Domestication of the scallops and the fishermen of St Brieuc Bay', in (ed.) J. Law (1986), *Power, Action and Belief: A New Sociology of Knowledge?*. London: Routledge, pp. 196-223.

Callon, M. (1986b), 'The sociology of an actor-network: The case of the electric vehicle', in (eds.) Callon, M., J. Law & A. Rip, *Mapping the Dynamics of Science and Technology: Sociology of Science in the Real World*, London: Macmillan, pp. 19-34.

Callon, M. & V. Rabeharisoa (2004), 'Gino's lesson on humanity: Genetics, mutual entanglements and the sociologist's role', *Economy and Society*, 33(1), pp. 1-27.

Cardamone, C., K. Schawinski, M. Sarzi, S. Bamford, N. Bennert, C. Urry, C. Lintott, W. Keel, J. Parejko, R. Nichol, D. Thomas, D. Andreescu, P. Murray, M. Raddick, A. Slosar, A. Szalay & J. Vandenbery (2009), 'Galaxy Zoo Green Peas: Discovery of a class of compact extremely star-forming galaxies', *Monthly Notices of the Royal Astronomical Society*, 399, pp. 1191-1207.

Carlisle, J., J. Feezell, K. Michaud, E. Smith & L. Smith (2009), 'The public's trust in scientific claims regarding offshore oil drilling', *Public Understanding of Science*, 19(5), pp. 514-527.

Carusi, A. & G. De Grandis (forthcoming), 'The ethical work that regulations will not do', *Information, Communication & Society*, pp.

Chilvers, J. (2008), 'Deliberating competence: Theoretical and practitioner perspectives on effective participatory appraisal practice', *Science, Technology & Human Values*, 33(2), pp. 155-185.

Christensen, C., T. Aina, & D. Stainforth (2005), 'The challenge of volunteer computing with lengthy climate model simulations', *Proceedings of the First International Conference on e-Science and Grid Computing*.

Clark, H. & S. Brennan (1991), 'Ground in communication', in (eds.) Resnik, L., J. Levine & S. Teasley, *Perspectives on Socially Shared Cognition*, Washington, DC: American Psychological Association, pp. 127-149.

Coen-Pirani, D., A. Leon & S. Lugauer (2008), *The effect of household appliances on female labor force participation: Evidence from Micro Data* mimeo.

Cole, J. & S. Cole (1973), *Social Stratification in Science*, Chicago: University of Chicago Press.

Coleman, E., G. Golub & A. Golub (2008), 'Hacker practice: moral genres and the cultural articulation of liberalism', *Anthropological Theory*, 8(3), pp. 255-277.

Collins, H. (1985), *Changing Order: Replication and Induction in Scientific Practice*, London, Sage.

Collins, H. (2007), 'A new programme of research?', *Studies in History and Philosophy of Science A*, 38(4), pp. 615-620.

Collins, H. & R. Evans (2002), 'The Third Wave of Science Studies: Studies of Expertise and Experience', *Social Studies of Science*, pp. 235-296.

Collins, H. & R. Evans (2007), *Rethinking Expertise*, Chicago, Illinois: Chicago University Press.

Collins, M. & S. Knight (eds.) (2007), 'Ensembles and probabilities: A new era in the prediction of climate change', *Philosophical Transactions of the Royal Society: A*, 365(1857), pp. 1957-1970.

Connor, M. & M. Siegrist (2010), 'Factors influencing people's acceptance of gene technology: The role of knowledge, health expectations, naturalness, and social trust', *Science Communication*, 32(4), pp. 514-538.

Conselice, C. (2006), 'The fundamental properties of galaxies and a new galaxy classification system', *Monthly Notices of the Royal Astronomical Society*, 373, pp. 1389-1408.

Correll, S. (1995), 'The ethnography of an electronic bar: The lesbian café', *Journal of Contemporary Ethnography*, 24(3), pp. 270-298.

Costa, F., L. Silva, I. Kelley & I. Taylor (2007), *Peer-to-peer techniques for data distribution in Desktop Grid computing platforms: CoreGRID Technical Report TR-0095*, available online at: http://www.coregrid.net/mambo/images/stories/TechnicalReports/tr-0095.pdf (accessed 22 December 2011).

Cvetkovich, G. & R. Lofstedt (1999), *Social Trust and the Management of Risk*, London: Earthscan.

Darg, D., S. Kaviraj, C. Lintott, K. Schawinski, M. Sarzi, S. Bamford, J. Silk, R. Proctor, D. Andreescu, P. Murray, R. Nichol, M. Raddick, A. Slosar, A. Szalay, D. Thomas & J. Vandenberg (2010), 'Galaxy Zoo: The fraction of merging galaxies in the SDSS and their morphologies', *Monthly Notices of the Royal Astronomical Society*, 401, pp. 1043-1058.

Darier, E., C. Gough, B. De Marchi, S. Funtowicz, R. Grove-White, D. Kitchener, A. Pereira, S. Shackley & B. Wynne (1999), 'Between democracy and expertise? Citizens' participation and environmental integrated assessment in Venice (Italy) and St Helens (UK)', *Journal of Environmental Policy and Planning*, 1(2), pp. 103-120.

Davies, G. & J. Burgess (2004), 'Challenging the 'view from nowhere': Citizen reflections on specialist expertise in a deliberative process', *Health & Place*, 10(4), pp. 349-361.

de Laet, M. & A. Mol (2000), 'The Zimbabwe bush pump: Mechanics of a fluid technology', *Social Studies of Science*, 30(2), pp. 225-263.

de Rooij, J., R. Verburg, E. Andriessen, & D. den Hartog (2007), 'Barriers for shared understanding in virtual teams: A leader perspective', *Electronic Journal for Virtual Organizations and Networks*, 9, pp. 64-77.

de Vaucouleurs, G., A. de Vaucouleurs, H. Corwin, R. Buta, G. Paturel & P. Fouque (1991), *Third Reference Catalogue of Bright Galaxies*, New York: Springer.

Delborne, J., A. Anderson, D. Kleinman, M. Colin & M. Powell (2011), 'Virtual deliberation? Prospects and challenges for integrating the Internet in consensus conferences', *Public Understanding of Science*, 20(3), pp. 367-384.

Deuten, J. (2003), *Cosmopolitanising Technologies: A Study of Four Emerging Technological Regimes*, Enschede: Twente University Press.

Devorkin, D. (1980), 'The maintenance of a scientific institution: Otto Struve, the Yerkes Observatory, and its optical bureau during the Second World War', *Minerva*, 18(4), pp. 595-623.

Dibbell, J. (1993), 'A Rape in Cyberspace', *The Village Voice*, 21 December 1993, available online at: http://www.juliandibbell.com/texts/bungle_vv.html (accessed 31 August 2011).

Dolby, R. (1971), 'Sociology of knowledge in social science', *Social Studies of Science*, 1, pp. 3-21.

Douglas, M. (1966), *Purity and Danger: An Analysis of Concepts of Pollution and Taboo*, London: Routledge & Kegan Paul.

Edelman, M. (1985), *The Symbolic Uses of Politics*, Chicago: University of Illinois Press.

Einsiedel, E. (1992), 'Framing science and technology in the Canadian press', *Public Understanding of Science*, 1(11), pp. 89-101.

Einsiedel, E., E. Jelsoe & T. Breck (2001), 'Publics at the technology table: The consensus conference in Denmark, Canada, and Australia', *Public Understanding of Science*, 10(1), pp. 83-98.

Epstein, S. (1995), 'The construction of lay expertise: AIDS activism and the forging of credibility in the reform of clinical trials', *Science, Technology & Human Values*, 20(4), pp. 408-437.

Epstein, S. (1996), *Impure Science: AIDS, Activism, and the Politics of Knowledge*, Berkeley: University of California Press.

Epstein, S. (2000), 'Democracy, expertise, and AIDS treatment activism', In (ed.) D. Kleinmann, *Science, Technology, and Democracy*, Albany: State University of New York Press, pp. 15-32.

Escobar, A. (1994), 'Welcome to Cyberia: Notes on the anthropology of cyberculture', *Current Anthropology*, 35(3), pp. 211-231.

Ess, C. & AoIR (Association of Internet Researchers) (2002), *Ethical Guidelines for Internet Research*, available from: http://www.aoir.org/reports/ethics.pdf (Retrieved 30 September 2011).

Ess, C. (2006), 'Ethical pluralism and global information ethics', *Ethics and Information Technology*, 8, pp. 215-226.

Felt, U., M. Fochler, A. Mueller & M. Strassnig (2009), 'Unruly ethics: On the difficulties of a bottom-up approach to ethics in the field of genomics', *Public Understanding of Science*, 18(3), pp. 357-371.

Finkbeiner, A. (2010), *A Grand and Bold Thing: An Extraordinary New Map of the Universe Ushering in a New Era of Discovery*, New York: Free Press.

Forman, P. (1971), 'Weimar culture, causality, and quantum theory, 1918-1927: Adaption by German physicists and mathematicians to a hostile intellectual environment', in *Historical Studies in the Physical Sciences, vol. 3*, (ed.) McCormmach, R., Philadelphia: University of Philadelphia Press, pp. 1-115.

Forman, P. (1987), 'Behind quantum electronics: National security as basis for physical research in the United States, 1940-1960', *Historical Studies in the Physical and Biological Sciences*, 18(1), pp. 149-229.

Fortson, L., K. Masters, R. Nichol, K. Borne, E. Edmondson, C. Lintott, J. Raddick, K. Schawinski, J. Wallin (2011), 'Galaxy Zoo: Morphological classification and citizen science', in (eds.) Way, M., J. Scargle, K. Ali & A. Srivastava, *Advances in Machine Learning and Data Mining for Astronomy*, London: Chapman & Hall.

Fox Keller, E. (1983), *A Feeling for the Organism: The Life and Work of Barbara McClintock*, New York: Freeman.

Fox Keller, E. (1985), *Reflections on Gender and Science*,

Fox Keller, E. (2011), 'What are climate scientists to do?', *Spontaneous Generations: A Journal for the History and Philosophy of Science*, 5(1), pp. 19-26.

French, S. & C. Bayley (2011), 'Public participation: Comparing approaches', *Journal of Risk Research*, 14(2), pp. 241-257.

Galison, P. (1997), *Image and Logic: A Material Culture of Microphysics*, Chicago: University of Chicago Press.

Galison, P. & B. Hevly (1992), *Big Science: The Growth of Large-Scale Research*, Stanford, California: Stanford University Press.

Galison, P. & D. Stump (1996), *The Disunity of Science: Boundaries, Contexts, and Power*, Stanford, California: Stanford University Press.

Garfinkel, H., M. Lynch & E. Livingston (1981), 'The work of a discovering science construed with materials from the optically discovered quasar', *Philosophy of Social Sciences*, 11, pp. 131-158.

Geertz, C. (1973), *The Interpretation of Cultures*, New York: Basic Books.

Gieryn, T. (1983), 'Boundary-work and the demarcation of science from non-science: Strains and interests in professional ideologies of scientists', *American Sociological Review*, 48, pp. 781-795.

Gieryn, T. (1995), 'Boundaries of science', in (eds.) Jasanoff, S., G. Markle, J. Petersen & T. Pinch (1995), *Handbook of Science and Technology Studies*, London: Sage, pp. 393-444.

Gieryn, T. (1999), *Cultural Boundaries of Science: Credibility on the Line*, Chicago: University of Chicago Press.

Gilbert, N. & M. Mulkay (1984), *Opening Pandora's Box: A Sociological Analysis of Scientists' Discourse*, Cambridge: Cambridge University Press.

Gillespie, B., D. Eva & R. Johnston (1979), 'Carcinogenic risk assessment in the United States and Great Britain: The case of Aldrin/Dieldrin', *Social Studies of Science*, 9(3), pp. 265-301.

Glaser, B. & A. Strauss (1967), *The Discovery of Grounded Theory*, Chicago, Ill.: Aldine.

Goldin, D. (2004), 'Bold missions and the big picture: The societal impact of America's space program', *Technology in Society*, 26, pp. 361-370.

Gregory, J. & S. Miller (1998), *Science in Public: Communication, Culture and Credibility*, New York: Plenum Press.

Gross, J. & B. Lewenstein (2005), *The Salience of Small: Nanotechnology Coverage in the American Press, 1986-2004*, Media and Society Research Group, Ithaca: Cornell University Press.

Gunn, J., W. Siegmund, E. Mannery, R. Owen, C. Hull, R. Leger, L. Carey, G. Knapp, D. York, W. Boroski, S. Kent, R. Lupton, C. Rockosi, M. Evans, P. Waddell, J. Anderson, J. Annis, J. Barentine, L. Bartoszek, S. Bastian, S. Bracker, H. Brewington, C. Briegel, J. Brinkmann, Y. Brown, M. Carr, P. Czarapata, C. Drennan, T. Dombeck, G. Federwitz, B. Gillespie, C. Gonzales, S. Hansen, M. Harvanek, J. Hayes, W. Jordan, E. Kinney, M. Klaene, S. Kleinman, R. Kron, J. Kresinski, G. Lee, S. Limmongkol, C. Lindenmeyer, D. Long, C. Loomis, P. McGehee, P. Mantsch, E. Neilsen Jr., R. Neswold, P. Newman, A. Nitta, J. Peoples Jr., J. Pier, P. Prieto, A. Prosapio, C. Rivetta, D. Schneider, S. Snedden & S. Wang (2006), 'The 2.5m telescope of the Sloan Digital Sky Survey', *The Astronomical Journal*, 131, pp. 2332-2359.

Guston, D. (1999), ‘Evaluating the first US consensus conference: The impact of the citizens’ panel on telecommunications and the future of democracy’, *Science, Technology & Human Values*, 24(4), pp. 451-482.

Hagstrom, W. (1965), *The Scientific Community*, New York: Basic Books.

Haklay, M. (2010), ‘ “Extreme” Citizen Science’, in *Proceedings of the London Citizen Cyberscience Project*, London, September 2010.

Hammersley, M. & P. Atkinson (1983), *Ethnography: Principles in to Practice*, London: Tavistock Publications.

Haraway, D. (1988), ‘Situated knowledges: The science question in feminism and the privilege of partial perspectives’, *Feminist Studies*, 14(3), pp. 575-599.

Haraway, D. (1991), *Simians, Cyborgs and Women: The Reinvention of Nature*.

Haraway, D. (2004), ‘A Manifesto for Cyborgs: Science, technology and socialist feminism in the 1980s’, in (ed.) Haraway, D., *The Haraway Reader*, New York: Routledge, pp. 7-46.

Hardwig, J. (1985), ‘Epistemic dependence’, *The Journal of Philosophy*, 82: 335-49.

Hardwig, J. (1991), ‘The role of trust in knowledge’, *The Journal of Philosophy*, 88: 693-708.

Harper, R. (1998), *Inside the IMF: An Ethnography of Documents, Technology and Organisational Action*, San Diego, CA: Academic Press.

Hartman, J. (1997), ‘The popularization of science through citizen volunteers’, *Public Understanding of Science*, 6, pp. 69-86.

Haythornthwaite, C. & B. Wellman (eds.) (2002), *The Internet in Everyday Life*, London: Blackwell.

Heath, D., E. Koch, B. Ley & M. Montoya (1999), ‘Nodes and queries: Linking locations in networked fields of inquiry’, *American Behavioral Scientist*, 43(3), pp. 450-463.

Hess, D. (2001), ‘Ethnography and the development of science and technology studies’, in (eds.) Atkinson, P., A. Coffey, S. Delamont, J. Lofland & L. Lofland, *Sage Handbook of Ethnography*, Thousand Oaks, California: Sage, pp. 234-245.

Hey, T., S. Tansley & K. Tolle (2009), *The Fourth Paradigm: Data-Intensive Scientific Discovery*, Redmond, WA: Microsoft Research.

Hilgartner, S. (1990), ‘The dominant view of popularization: Conceptual problems, political uses’, *Social Studies of Science*, 20, pp. 519-539.

Hinchliffe, S. (2001), ‘Indeterminacy in-decisions - science, policy and politics in the BSE (Bovine Spongiform Encephalopathy) crisis’, *Transactions - Institute of British Geographers,* 26(2), pp. 182-204.

Hinds, P. & S. Weisband (2003), ‘Knowledge sharing and shared understanding in virtual teams’, in (eds.) Gibson, C. & S. Cohen, *Virtual Teams That Work: Creating Conditions for Virtual Team Effectiveness*, San Francisco, CA: Jossey-Bass, pp. 21-36.

Hine, C. (2000), *Virtual Ethnography*, London: Sage.

Hine, C. (2007), ‘Connective ethnography for the exploration of e-Science’, *Journal of Computer-Mediated Communication*, 12(2), article 14, available at: http://jcmc.indiana.edu/vol12/issue2/hine.html (accessed 31 August 2011).

Ho, S., D. Scheufele & E. Corley (2010), ‘Value predispositions, mass media, and attitudes toward nanotechnology: The interplay of public and experts’, *Science Communication*, 32(3), pp. 1-34.

Hobbs, D. (1988), *Doing the Business: Entrepreneurship, the Working Class and Detectives in the East End of London*, Oxford: Clarendon Press.

Holohan, A. & A. Garg (2005), 'Collaboration online: The example of Distributed Computing', *Journal of Computer-Mediated Communication*, 10(4), available online at: jcmc.indiana.edu/vol10/issue4/holohan.html (accessed 22 December 2011).

House of Lords (2000), *Science and Society*, London: The Stationary Office.

Hubble, E. (1926), 'Extra-galactic nebulae', *Astrophysics Journal*, 64, pp. 321-369.

Hutchby, I. (2001), 'Technologies, texts and affordances', *Sociology*, 35(2), pp. 441-456.

Irwin, A. (1995), *Citizen Science: A Study of People, Expertise and Sustainable Development (Environment & Society)*, Routledge: London.

Irwin, A., A. Dale & D. Smith (1996), 'Science and Hell's kitchen: The local understanding of hazard issues', In (eds.) A. Irwin & B. Wynne, *Misunderstanding Science?: The public reconstruction of science and technology*, Cambridge University Press: Cambridge, pp. 47- 64.

Jasanoff, S. (1997), 'Civilization and madness: The great BSE scare of 1996', *Public Understanding of Science*, 6, pp. 221-232.

Jasanoff, S. (2007), *Designs on Nature: Science and Democracy in Europe and the United States*, Princeton, NJ: Princeton University Press.

Jones, R. (1994), 'The ethics of research in cyberspace', *Internet Research*, 4(3), pp. 30-35.

Joss, S. (1995), 'Evaluating consensus conferences: Necessity or luxury', In (eds.) S. Joss & J. Durant, *Public Participation in Science: The Role of Consensus Conferences in Europe*, London: The Science Museum, pp. 89-108.

Kahan, D., H. Jenkins-Smith & D. Braman (2011), 'Cultural cognition of scientific consensus', *Journal of Risk Research*, 14(2), pp. 147-174.

Kaminstein, D. (1996), 'Persuasion in a toxic community: Rhetorical aspects of public meetings', *Human Organization*, 55(4), pp. 458-464.

Kathlene, L. & J. Martin (1991), 'Enhancing citizen participation: Panel designs, perspectives, and planning', *Journal of Policy Analysis and Management*, 10, pp. 46-63.

Kerr, A., S. Cunningham-Burley & R. Tutton (2007), 'Shifting subject positions', *Social Studies of Science*, 37(3), pp. 385-411.

King, M. (1971), 'Reason, tradition, and the progressiveness of science', *History and Theory*, 6, pp. 3-22.

Kling, R. (1992), 'When gunfire shatters bone: Reducing sociotechnical systems to social relationships', *Science, Technology and Human Values*, 17(3), pp. 381-385.

Knorr-Cetina, K. (1981), *The Manufacture of Knowledge: An Essay on the Constructivist and Contextual Nature of Science*, Oxford and New York: Pergamon.

Knorr-Cetina, K. (1999), *Epistemic Cultures: How the Sciences Make Knowledge*, Cambridge, Mass.: Harvard University Press.

Kondo, D., D. Anderson & J. McLeod (2007), 'Performance evaluation of scheduling policies for Volunteer Computing', *IEEE International Conference on e-Science and Grid Computing*, Bangalore, India, December 2007, pp. 415-422.

Korpela, E., D. Werthimer, D. Anderson, J. Cobb & M. Lebofsky (2001), 'SETI@Home – massively distributed computing for SETI', *Computing in Science & Engineering*, 3(1), pp. 78-83.

Kragh, H. (1980), 'Anatomy of a priority dispute: The case of element 72', *Centaurus*, 23(4), pp. 275-301.

Kragh, H. (1996), *Cosmology and Controversy: The Historical Development of Two Theories of the Universe*, Princeton: Princeton University Press.

Kuhn, T. S. (1962), *The Structure of Scientific Revolutions*, Chicago: University of Chicago Press.

Lambert, H. & H. Rose (1996), ‘Disembodied knowledge? Making sense of medical science’, In (eds.) A. Irwin & B. Wynne, *Misunderstanding Science?: The Public Reconstruction of Science and Technology*, Cambridge University Press: Cambridge, pp. 65-83.

Land, K., A. Slosar, C. Lintott, D. Andreescu, S. Bamford, P. Murray, R. Nichol, M. Raddick, K. Schawinski, A. Szalay, D. Thomas & J. Vandenberg (2008), ‘Galaxy Zoo: The large-scale spin statistics of spiral galaxies in the Sloan Digital Sky Survey’, *Monthly Notices of the Royal Astronomical Society*, 388, pp. 1686-1693.

Lankford, J. (1981), ‘Amateurs versus professionals: The controversy over telescope size in late Victorian science’, *Isis*, 72(1), pp. 11-28.

Latour, B. (1987), *Science in Action: How to Follow Scientists and Engineers through Society*. Cambridge, Mass.: Harvard University Press.

Latour, B. (1988), *The Pasteurization of France*, Cambridge, Mass.: Harvard University Press.

Latour, B. (2007), *Reassembling the Social: An Introduction to Actor-Network-Theory*, Oxford, Oxford University Press.

Latour, B. & S. Woolgar (1979), *Laboratory Life: The Construction of Scientific Facts*, Princeton: Princeton UP.

Leach, J., D. Nafus & B. Krieger (2009), ‘Freedom imagined: Morality and aesthetics in Open Source Software design’, *Ethnos*, 74(1), pp. 51-71.

Lee, G. (1995), ‘A consensus conference from the point of view of a lay-panel member’, in (eds.) S. Joss & J. Durant, *Public Participation in Science: The role of consensus conferences in Europe*, London: The Science Museum, pp. 81-88.

Lee, R. & T. Garvin (2003), ‘Moving from information transfer to information exchange in health and health care’, *Social Science & Medicine*, 56, pp. 449-464.

Lengwiler, M. (2008), ‘Participatory approaches in science and technology: Historical origins and current practices in critical perspective’, *Science, Technology & Human Values*, 33(2), pp. 186-200.

Lenhard, J. & E. Winsberg (2010), ‘Holism, entrenchment, and the future of climate model pluralism’, *Studies in History and Philosophy of Modern Physics*, 41, pp. 253-262.

Levy, S. (1984), *Hackers: Heroes of the Computer Revolution*, New York: Doubleday.

Lewenstein, B. (1995), ‘Science and the media’, in (eds.) S. Jasanoff, G. Markle, J. Petersen & T. Pinch, *Handbook of Science and Technology Studies*, London: SAGE Publications, pp. 343-360.

Lien, M. (1987), *Marketing and Modernity*, Oxford: Berg.

Lindenschmidt, J. (2000), ‘A barnraising in cyberspace: Linux and the Free Software Movement’, *The Maine Scholar*, 13, pp. 123-138.

Lintott, C. (2010), *Learning from Galaxy Zoo*, available at: http://projects.oucs.ox.ac.uk/runcoco/downloads/Chris_Lintott_RunCoCo.pdf (accessed 25 October 2011).

Lintott C., K. Schawinski, A. Slosar, K. Land, S. Bamford, D. Thomas, M. Raddick, R. Nichol, A. Szalay, D. Andreescu, P. Murray & J. Vandenberg (2008), ‘Galaxy Zoo: Morphologies derived from visual inspection of galaxies from the Sloan Digital Sky Survey’, *Monthly Notices of the Royal Astronomical Society*, 389, pp. 1179-1189.

Lintott, C., K. Schawinski, W. Keel, H. van Arkel, N. Bennert, E. Edmondson, D. Thomas, D. Smith, P. Herbert, M. Jarvis, S. Virani, D. Andreescu, S. Bamford, K. Land, P. Murray, R. Nichol, J. Raddick, A. Slosar, A. Szalay & J. Vandenbery (2009), ‘Galaxy Zoo: Hanny’s Voorwerp, a quasar light echo?’, *Monthly Notices of the Royal Astronomical Society*, 399, pp. 129-141.

Lintott, C., K. Schawinski, S. Bamford, A. Slosar, K. Land, D. Thomas, E. Edmondson, K. Masters, R. Nichol, J. Raddick, A. Szalay, D. Andreescu, P. Murray & J. Vandenberg (2011), 'Galaxy Zoo 1: Data release of morphological classifications for nearly 900,000 galaxies', *Monthly Notices of the Royal Astronomical Society*, 410, pp. 166-179.

LSST Science Collaborations (2009), *LSST Science Book*, available from: http://www.lsst.org/lsst/scibook (accessed 30 September 2011).

Lynch, M. (1985), *Art and Artifact in Laboratory Science: A Study of Shop Work and Shop Talk in a Research Laboratory*, London: Routledge & Kegan Paul.

Macintyre, S. (1995), 'The public understanding of science or the scientific understanding of the public? A review of the social context of the "new geneticist"', *Public Understanding of Science*, 4(3), pp. 223-232.

MacNaghten, P., M. Kearnes & B. Wynne (2005), 'Nanotechnology, Governance, and Public Deliberation', *Science Communication*, 27, pp. 268-291.

Maddox, B. (2002), *Rosalind Franklin: The Dark Lady of DNA*, London: HarperCollins.

Malinowski, B. (1922), *Argonauts of the Western Pacific*, London: Routledge & Kegan Paul.

Markham, A. (1998), *Life Online: Researching Real Experience in Virtual Space*, Walnut Creek, CA: Altamira.

Markham, A. (2006), 'Method as ethic, ethic as method: A case for reflexivity in qualitative ICT research', *Journal of Information Ethics*, 15(2), pp. 37-55.

Markham, A. (2011), 'Internet Research', in (ed.) Silverman, D., *Qualitative Research: Theory, Method and Practices*, London: Sage, pp. 111-128.

Markham, A. & N. Baym (eds.) (2009), *Internet Inquiry: Conversations About Method*, Thousand Oaks, CA: Sage.

Martin, E. (1998), 'Anthropology and the cultural study of science', *Science Technology & Human Values*, 23(1), pp. 24-44.

Massey, N., T. Aina, M. Allen, C. Christensen, D. Frame, D. Goodman, J. Kettleborough, A. Martin, S. Pascoe & D. Stainforth (2006), 'Data access and analysis with distributed federated data servers in climateprediction.net', *Advances in Geosciences*, 8, pp. 49-56.

Masters, K., R. Nichol, S. Bamford, M. Mosleh, C. Lintott, D. Andreescu, E. Edmondson, W. Keel, P. Murray, M. Raddick, K. Schawinski, A. Slosar, A. Szalay, D. Thomas & J. Vandenberg (2010), 'Galaxy Zoo: Dust in spirals', *Monthly Notices of the Royal Astronomical Society*, 404, pp. 792-813.

Masters, K., R. Nichol, B. Hoyle, C. Lintott, S. Bamford, E. Edmondson, L. Fortson, W. Keel, K. Schawinski, A. Smith & D. Thomas (2011), 'Galaxy Zoo: Bars in disk galaxies', *Monthly Notices of the Royal Astronomical Society*, 411, pp. 2026-2035.

McCormick, S., P. Brown & S. Zavestoski (2003), 'The personal is scientific, the scientific is political: The public paradigm of the environmental breast cancer movement', *Sociological Forum*, 18(4), pp. 545-576.

McSherry, C. (2003), 'Uncommon controversies: Legal mediations of gift and market models of authorship', in (eds.) Biagioli, M. & P. Galison (2003), *Scientific Authorship: Credit and Intellectual Property in Science*, New York: Routledge, pp. 225-250.

Mead, M. (1928), *Coming of Age in Samoa: A Psychological Study of Primitive Youth for Western Civilization*, New York: W. Morrow.

Merton, R. (1949), 'The normative structure of science', in (ed.) Merton, R. (1973), *The Sociology of Science: Theoretical and Empirical Investigations*, Chicago: University of Chicago Press, pp. 268-280.

Michael, M. (1992), 'Lay discourses of science: Science-in-general, science-in-particular, and self', *Science, Technology & Human Values*, 17(3), pp. 313-333.

Miller, D. & D. Slater (2000), *The Internet: An Ethnographic Approach*, London: Berg.

Miller, S. (2001), 'Public understanding of science at the crossroads', *Public Understanding of Science*, 10(1), pp. 115-120.

Mitroff, I. (1974), 'Norms and counter-norms in a select group of the Apollo moon scientists: A case study of the ambivalence of scientists', *American Sociological Review*, 39(4), pp. 579-595.

Mol, A. (2002), *The Body Multiple: Ontology in Medical Practice*, Durham, North Carolina: Duke University Press.

Mol, A. & J. Law (1994), 'Regions, networks and fluids: Anaemia and social topology', *Social Studies of Science*, 24(4), pp. 641-671.

Morris, S. (2002), 'First person shooters: A game apparatus', in (eds.), King, G. & T. Krzywinska, *ScreenPlay: Cinema/videogames/interfaces*, London: Wallflower Press, pp. 81-97.

Moseley, R. (1977), 'Tadpoles and frogs: Some aspects of the professionalization of British Physics, 1870-1939', *Social Studies of Science*, 7(4), pp. 423-446.

National Research Council (1991), *The Decade of Discovery in Astronomy and Astrophysics*, Washington, D.C.: National Academy Press.

Natural Environment Research Council (2007), *Final Report to the Natural Environment Research Council on the Coupled Ocean Atmosphere Process and European Climate (COAPEC) Directed Programme*, available online at: http://www.nerc.ac.uk/research/programmes/coapec/documents/coapecfinalreportv2pt1.pdf (accessed 11 November 2011).

Nelkin, D. (1979), 'Scientific knowledge, public policy, and democracy: A review essay', *Science Communication*, 1(1), pp. 106-122.

Nelkin, D. (1994), 'Promotional metaphors and their popular appeal', *Public Understanding of Science*, 3(1), pp. 25-31.

Nelkin, D. (1995), *Selling Science: How the Press Covers Science and Technology*, New York: W. H. Freeman and Company.

Nelkin, D. & S. Gilman (1988), 'Placing blame for devastating disease', *Social Research*, pp. 361-378.

Nerlich, B. (2010), ' "Climategate": Paradoxical metaphors and political paralysis', *Environmental Values*, 19(4), pp. 419-442.

Nisbet, M. & B. Lewenstein (2002), 'Biotechnology and the American media: The policy process and the elite press, 1970 to 1999', *Science Communication*, 23(4):359–91.

Noble, D. (1977), *America By Design*, New York: Alfred A. Knopf.

Nov, O., D. Anderson & O. Arazy (2010), 'Volunteer Computing: A model of the factors determining contribution to community-based scientific research', *International World Wide Web Conference (IW3)*, Raleigh, North Carolina, April 2010, pp. 741-750.

Nov, O., O. Arazy & D. Anderson (2011), 'Technology-mediated citizen science participation: A motivational model', *Proceedings of the AAAI International Conference on Weblogs and Social Media (ICWSM 2011)*, Barcelona, Spain, July 2011.

Nowotny, H. (1979), 'Experts in a participatory experiment: The Austrian debate on nuclear energy', In (ed.) H. Skoie, *Scientific Expertise and the Public*, Oslo: Studies in Research and Higher Education.

Nowotny, H. (1981), 'Experts and their expertise: On the changing relationship between experts and their public', *Bulletin of Science, Technology & Society*, 1, pp. 235-241.

Nowotny, H. (1993), ‘Socially distributed knowledge: Five spaces for science to meet the public’, *Public Understanding of Science*, 2(4), pp. 307-319.

O’Donoghue, T. & K. Punch (2003), *Qualitative Educational Research in Action: Doing and Reflecting*, London: RoutledgeFalmer.

Olson, G., N. Bos, N. & A. Zimmerman (eds.) (2008), *Scientific Collaboration on the Internet*, Boston, Massachusetts: MIT Press.

Olson, G. & J. Olson (2000), ‘Distance matters’, *Human Computer Interaction*, 15, pp. 139-179.

Orr, J. (1996), *Talking about Machines: An Ethnography of a Modern Job*, Ithaca, NY: ILR Press.

Parker, W. (2010), ‘Predicting weather and climate: Uncertainty, ensembles and probability’, *Studies in History and Philosophy of Science*, 41, pp. 263-272.

Parry, S. (2009), ‘Stem cell scientists’ discursive strategies for cognitive authority’, *Science as Culture*, 18(1), pp. 89-114.

Parthasarathy, S. (2010), ‘Breaking the expertise barrier: Understanding activist strategies in science and technology policy domains’, *Science and Public Policy*, 37(5), pp. 355-367.

Petts, J. (1995), ‘Waste management strategy development: A case study of community involvement and consensus-building in Hampshire’, *Journal of Environmental Planning and Management*, 38(4), pp. 519-536.

Piani, C., D. Frame, D. Stainforth & M. Allen (2005), ‘Constraints on climate change from a multi-thousand member ensemble of simulations’, *Geophysical Review Letters*, 32, L23825.

Pickering, A. (1995), *The Mangle of Practice: Time, Agency, and Science*, Chicago: University of Chicago Press.

Pickering, A. (2006), ‘Concepts and the Mangle of Practice: Constructing quaternions’, in R. Hersh (ed.), *18 Unconventional Essays on the Nature of Mathematics*, New York: Springer, pp. 250-288.

Pinch, T. & W. Bijker (1984), ‘The social construction of facts and artefacts: Or how the sociology of science and the sociology of technology might benefit each other’, *Social Studies of Science*, 14(3), pp. 399-441.

Plumlee, J., J. Starling & K. Kramer (1985), ‘Citizen participation in water quality planning’, *Administration and Society*, 16(4), pp. 455-473.

Polanyi, M. (1962), ‘The Republic of Science’, *Minerva*, 1(3), pp. 54-73.

Popper, K. (1959, *The Logic of Scientific Discovery*, London: Routledge.

Porter, T. (1996), *Trust in Numbers: The Pursuit of Objectivity in Science and Public Life*, Princeton: Princeton University Press.

Portolano, M. (2000), ‘John Quincy Adams’s rhetorical crusade for astronomy’, *Isis*, 91(3), pp. 480-503.

Potter, J. & M. Wetherell (1994), ‘Analyzing discourse’, in (eds.) Bryman, A. & R. Burgess, *Analyzing Qualitative Data*, London: Routledge, pp. 47-66.

Pourebrahimi, B., K. Bertels & S. Vassiliadis (2005), ‘A survey of peer-to-peer networks’, *The 16th Annual Workshop on Circuits, Systems and Signal Processing*.

Powell, M., M. Colin, D. Kleinman, J. Delborne & A. Anderson (2011), ‘Imagining ordinary citizens? Conceptualized and actual participants for deliberations on emerging technologies’, *Science as Culture*, 20(1), pp. 37-70.

Prasad, A. (2007), ‘The (amorphous) anatomy of an invention: The case of Magnetic Resonance Imaging (MRI)’, *Social Studies of Science*, 37(4), pp. 533-560.

Price, D. (1962), *Government and Science*, Oxford: Oxford University Press.

Price, D. (1965), *The Scientific Estate*, Cambridge, Mass.: Harvard University Press.

Raddick, J., G. Bracey, P. Gay, C. Lintott, P. Murray, K. Schawinski, A. Szalay & J. Vandenberg (2010), 'Galaxy Zoo: Exploring the motivations of citizen science', *Astronomy Education Review*, 9(1), available online at: aer.aas.org/resources/1/aerscz/v9/i1/p010103_s1 (accessed 22 December 2011).

Ravetz, J. (1971), *Scientific Knowledge and its Problems*, Oxford: Oxford University Press.

Rezabek, L. & J. Cochenour (1998), 'Visual cues in Computer-Mediated Communication: Supplementing text with emoticons', *Journal of Visual Literacy,* 18(2), pp. 201-215.

Roth, W.-M., J. Riecken, L. Pozzer-Ardenghi, R. McMillan, B. Storr, D. Tait, G. Bradshaw & T. Pauluth Penner (2004), 'Those who get hurt aren't always being heard: Scientist-resident interactions over community water', *Science, Technology & Human Values*, 29, pp. 153- 183.

Rowe, G. & L. Frewer (2004), 'Evaluating public-participation exercises: A research agenda', *Science, Technology & Human Values*, 29(4), pp. 512-56.

Roy, R. (1981), 'The reformation of science', *Bulletin of Science, Technology & Society*, 1, pp. 3-9.

Schawinski, K., S. Kaviraj, S. Khochfar, S.-J. Yoon, S. K. Yi, J.-M. Deharveng, A. Boselli, T. Barlow, T. Conrow, K. Forster, P. Friedman, D. C. Martin, P. Morrissey, S. Neff, D. Schiminovich, M. Seibert, T. Small, T. Wyder, L. Bianchi, J. Donas, T. Heckman, Y.-W. Lee, B. Madore, B. Milliard, R. M. Rich & A. Szalay (2007), 'The effect of environment on the ultraviolet color-magnitude relation of early-type galaxies', *Astrophysics Journal Supplement Series*, 173, pp. 512-523.

Schawinski, K., C. Urry, S. Virani, P. Coppi, S. Bamford, E. Treister, C. Lintott, M. Sarzi, W. Keel, S. Kaviraj, C. Cardamone, K. Masters, N. Ross, D. Andreescu, P. Murray, R. Nichol, M. Raddick, A. Slosar, A. Szalay, D. Thomas & J. Vandenberg (2010), 'Galaxy Zoo: The fundamentally different co-evolution of supermassive black holes and their early- and late-type host galaxies', *Astrophysics Journal*, 711, pp. 284-305.

Schilhab, T. (2007), 'Interactional expertise through the looking glass: A peek at mirror neurons', *Studies in History and Philosophy of Science A*, 38(4), pp. 741-747.

Shackley, S., J. Risbey, P. Stone & B. Wynne (1999), 'Adjusting to policy expectations in climate change science: An interdisciplinary study of flux adjustments in coupled atmosphere ocean general circulation models', *Climactic Change*, 43, pp. 413-454.

Shapin, S. (1979), 'Homo phrenologicus: Anthropological perspectives on an historical problem', in B. Barnes & S. Shapin (eds.), *Natural Order: Historical Studies in Scientific Culture*, Sage, Beverly Hills, pp. 41-71.

Shapin, S. (1990), 'Science and the public', In (eds.) R. Olby, G. Cantor, J. Christie & M. Hodge, *Companion to the History of Modern Science*, pp. 990-1007.

Shapin, S. (1994), *A Social History of Truth: Gentility, Credibility, and Scientific Knowledge in Seventeenth-Century England*, Chicago and London: University of Chicago Press.

Slosar, A., K. Land, S. Bamford, C. Lintott, D. Andreescu, P. Murray, R. Nichol, M. Raddick, K. Schawinski, A. Szalay, D. Thomas & J. Vandenberg (2008), 'Galaxy Zoo: Chiral correlation function of galaxy spins', *Monthly Notices of the Royal Astronomical Society*, 392, pp. 1225-1233.

Smith, L., J. Smith, K. Arcand, R. Smith, J. Bookbinder & K. Keach (2010), 'Aesthetics & astronomy: Studying the public's perception and understanding of imagery from space', *Science Communication*, pp. 1-38.

Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marquis, K. Avery, M. Tignor & H. Miller (eds.) (2007), *Climate Change 2007: The Physical Science Basis, Contribution of Working Group 1 to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, New York: Cambridge University Press.

Spencer, B., R. Butler, K. Ricker, D. Marcusiu, T. Finholt, I. Foster, C. Kesselman & J. Birnholtz (2008), 'NEESgrid: Lessons learned for future cyberinfrastructure development', in (eds.) Olson, G., N. Bos & A.

Stainforth, D., J. Kettleborough, A. Martin, A. Simpson, R. Gillis, A. Akkas, R. Gault, M. Collins, D. Gavaghan, & M. Allen (2002), 'Climateprediction.net: Design principles for public-resource modeling research', *Proceedings of the 14th IASTED International Conference: Parallel and Distributed Computing and Systems*, pp. 32-38.

Stainforth, D., T. Aina, C. Christensen, M. Collins, N. Faull, D. J. Frame, J. Kettleborough, S. Knight, A. Martin, J. Murphy, C. Piani, D. Sexton, L. A. Smith, R. Spicer, A. Thorpe & M. Allen (2005), 'Uncertainty in predictions of the climate response to rising levels of greenhouse gases', *Nature*, 433, pp.403-406.

Stein, D. (1985), *Ada: A Life and a Legacy*, Cambridge, Massachusetts: Harvard University Press.

Stilgoe, J. (2007), 'The co-production of public uncertainty: UK scientific advice on mobile phone health risks', *Public Understanding of Science*, 16, pp. 45-61.

Strathern, M. (1992), *After Nature*, Cambridge: Cambridge University Press.

Thorne, B. (1980), ' "You still takin' notes?": Fieldwork and the problems of informed consent', *Social Problems*, 27(3), pp. 284-297.

Timmermans, S. (2003), 'A black technician and blue babies', *Social Studies of Science*, 33(2), pp. 197-229.

Traweek, S. (1988), *Beamtimes and Lifetimes: The World of High Energy Physicists*, Cambridge, Mass. and London: Harvard University Press.

Turkle, S. (1995), *Life on the Screen: Identity in the Age of the Internet*, New York: Simon and Schuster.

Varnhagen, C., M. Gushta, J. Daniels, T. Peters, N. Parmar, D. Law, R. Hirsch, B. Sadler Takach & T. Johnson (2005), 'How informed is online informed consent?', *Ethics & Behavior*, 15(1), pp. 37-48.

Vetter, J. (2011), 'Introduction: Lay participation in the history of scientific observation', *Science in Context*, 24(2), pp. 127-141.

Wacquant, L. (2008), *Urban Outcasts: A Comparative Sociology of Advanced Marginality*, Cambridge, MA: Polity Press.

Walls, J., G. Rowe & L. Frewer (2011), 'Stakeholder engagement in food risk management: Evaluation of an iterated workshop approach', *Public Understanding of Science*, 20(2), pp. 241-260.

Weart, S. (2010), *The Discovery of Global Warming*, Cambridge, Massachusetts: Harvard University Press.

Weber, M. (1946), *From Max Weber: Essays in Sociology*, trans. H. Gerth & C. Wright Mills, Oxford University Press, New York.

Weber, M. (1947), *Theory of Social and Economic Organization*, trans. G. Roth & C. Wittich, Oxford University Press, New York.

Webler, T., & S. Tuler (2000), 'Fairness and competence in citizen participation: Theoretical reflections from a case study', *Administration & Society*, 32(5), pp. 566-595.

Weinel, M. (2007), 'Primary source knowledge and technical decision-making: Mbeki and the AZT debate', *Studies in History and Philosophy of Science A*, 38(4), pp. 748-760.

Weingart, P., A. Engels & P. Pansegrau (2000), 'Risks of communication: Discourses on climate change in science, politics, and the mass media', *Public Understanding of Science*, 9(3), pp. 261-283.

Welsh, E., M. Jirotka & D. Gavaghan (2006), 'Post-genomic science: Cross-disciplinary and large-scale collaborative research and its organizational and technological challenges for the scientific research process', *Philosophical Transactions of the Royal Society A*, 364, pp. 1533-1549.

Wynne, B. (1982), *Rationality and Ritual: The Windscale Inquiry and Nuclear Decisions in Britain*, British Society for the History of Science: London.

Wynne, B. (1991), 'Knowledges in context', *Science, Technology & Human Values* 16(1), pp. 111-122.

Wynne, B. (1992), 'Risk and social learning', In (eds.) S. Krimsky & D. Golding, *Social Theories of Risk*, Westport, Connecticut: Praeger, pp. 279-297.

Wynne, B. (1993), 'Public uptake of science: A case for institutional reflexivity', *Public Understanding of Science*, 2, pp. 321-327.

Wynne, B. (1996a), 'May the sheep safely graze?', In (eds.) S. Lash, B. Szerszyski & B. Wynne (1996), *Risk, Environment and Modernity*, London: Sage Publications , pp. 44-83.

Wynne, B. (1996b), 'Misunderstood misunderstandings: Social identities and public uptake of science', In (eds.) A. Irwin & B. Wynne, *Misunderstanding Science?: The Public Reconstruction of Science and Technology*, Cambridge University Press: Cambridge, pp. 19- 46.

Wynne, B. (2001), 'Creating public alienation: Expert cultures of risk and ethics on GMOs', *Science as Culture*, 10(4), pp. 445-481.

Wynne, B. (2003), 'Seasick on the Third Wave? Subverting the hegemony of propositionalism: Response to Collins & Evans (2002)', *Social Studies of Science*, 33(3), pp. 401-417.

Wynne, B. (2006), 'Public engagement as a means of restoring public trust in science: Hitting the notes, but missing the music', *Community Genetics*, 9(3), pp. 211-220.

Wynne, B. & U. Felt, (2007), *Taking European Knowledge Society Seriously: Report of the Expert Group on Science and Governance, Science Economy and Society Directorate EUR 22700*, Brussels: Commission of the European Communities.

Yearley, S. (2000), 'Making systematic sense of public discontents with expert knowledge: Two analytical approaches and a case study', *Public Understanding of Science*, 9, pp. 105-122.

Zimmerman, *Scientific Collaboration on the Internet*, Boston, Massachusetts: MIT Press.