

Syracuse University

SURFACE

School of Information Studies - Dissertations

School of Information Studies (iSchool)

5-2012

Crowdsourcing Scientific Work: A Comparative Study of Technologies, Processes, and Outcomes in Citizen Science

Andrea Wiggins
Syracuse University

Follow this and additional works at: https://surface.syr.edu/it_etd



Part of the [Library and Information Science Commons](#)

Recommended Citation

Wiggins, Andrea, "Crowdsourcing Scientific Work: A Comparative Study of Technologies, Processes, and Outcomes in Citizen Science" (2012). *School of Information Studies - Dissertations*. 72.

https://surface.syr.edu/it_etd/72

This Dissertation is brought to you for free and open access by the School of Information Studies (iSchool) at SURFACE. It has been accepted for inclusion in School of Information Studies - Dissertations by an authorized administrator of SURFACE. For more information, please contact surface@syr.edu.

ABSTRACT

Crowdsourcing Scientific Work:
A Comparative Study of Technologies, Processes, and Outcomes in Citizen Science

by
Andrea Wiggins

Citizen science projects involve the public with scientists in collaborative research. Information and communication technologies for citizen science can enable massive virtual collaborations based on voluntary contributions by diverse participants. As the popularity of citizen science increases, scientists need a more thorough understanding of how project design and implementation decisions affect scientific outcomes.

Applying a comparative case study methodology, the study investigated project organizers' perspectives and experiences in Mountain Watch, the Great Sunflower Project, and eBird, three observation-based ecological citizen science projects in different scientific domains. Five themes are highlighted in the findings: the influence of project design approaches that favor science versus lifestyle; project design and organizing implications of engaging communities of practice; relationships between physical environment, technologies, participant experiences, and data quality; the constraints and affordances of information and communication technologies; and the relationship of resources and sustainability to institutions and scale of participation.

This research contributes an empirically-grounded theoretical model of citizen science projects, with comparative analysis that produced new insights into the design of technologies and processes to support public participation in the production of scientific knowledge.

CROWDSOURCING SCIENTIFIC WORK

*A Comparative Study of
Technologies, Processes, and Outcomes in Citizen Science*

Andrea Wiggins

B.A., Alma College, 2000

M.S.I., University of Michigan, 2007

Dissertation submitted to the Graduate School of Syracuse University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy, Information Science & Technology

May 2012



Andrea Wiggins 2012
Some Rights Reserved

Released under Creative Commons Attribution, Non-Commercial 3.0 License
<http://creativecommons.org/licenses/by-nc/us/>

Attribution can be accomplished through standard academic citation.

ACKNOWLEDGEMENTS

No dissertation is accomplished alone. I gratefully thank the many colleagues and friends helped bring this work to fruition, but a number of people deserve special acknowledgments.

My sincerest thanks to Kevin Crowston, my advisor, for his patient mentorship throughout my PhD studies and continual support in many forms. Kevin taught me the craft of research, unquestionably one of the finest gifts I have ever received. I could not have asked for a better role model and advisor.

My committee members also deserve thanks; it is a privilege and a pleasure to have the support of such fine scholars. Steve Sawyer provided invaluable opportunities for my scholarly development. Jian Qin fostered my interest in scientific collaboration. Rick Bonney shared access and unparalleled insight into the citizen science community. In addition to serving as my external reader, Geof Bowker was a truly inspirational mentor. Murali Venkatesh provided valuable feedback as my internal reader. John Burdick officiated as dissertation chair, but also encouraged me to engage new theoretical perspectives.

I owe an enormous debt of gratitude to the staff and partners of the citizen science projects that served as case studies. The Cornell Lab of Ornithology and eBird staff took my feedback to heart and encouraged my interest in birding, an unexpected pleasure that has truly enriched my life. The staff of the Great Sunflower Project offered much more than time, insight, and the occasional interview. The staff of the Appalachian Mountain Club brought greater depth to my research, and the Mountain Watch experience helped me better appreciate my personal strengths.

The larger citizen science community has been incredibly generous and supportive, particularly the individuals and organizations that supported the early stages of this research: Abe Miller-Rushing, Jake Weltzin, Geri Tierney, Northeast Regional Phenology Network, USA National Phenology Network, Acadia National Park, Boston Harbor Islands National Historical Park, and the National Park Service Northeast Temperate Network Inventory and Monitoring Program. The members of the DataONE Public Participation in Scientific Research Working Group gave encouragement and mentorship that I never expected but deeply appreciate: Rick Bonney, Eric Graham, Sandra Henderson, Gretchen LeBuhn, Kelly Lotts, Greg Newman, Eric Russell, Rob Stevenson, Jake Weltzin, and of course, Bill Michener. Jennifer Shirk helped me in too many ways to enumerate; she is not only a colleague, but a dear friend.

I thank the mentors in the academic community who freely shared advice and insight: Elizabeth Davidson, Paul Dourish, Anatoliy Gruzd, Christine Halverson, Rob Kauffman, Sara Kiesler, Wayne Lutters, David MacDonald, Michael Muller, Carsten Oesterlund, Olivia Sheng, Diane Sonnenwald, and Michael Twidale.

The quality of this dissertation is due in no small part to the devoted efforts of my writing group, to whom I am truly grateful. Jaime Snyder and Mohammad Jarrahi provided unflagging support and assistance at every step of the process. Their involvement made my work better in every way. I also thank all of my PhD colleagues for their comradeship and support, especially Isabelle Fagnot, Mary Grace Flaherty, Joe Rubleske, and Sarah Webb. Extra special thanks go to my “academic big brother”, James Howison, for his patience, compassion, wisdom, and humor.

Thanks are due to the Michiganders for their constant encouragement: Chris Grant, Peter Honeyman, Mick McQuaid, and Scott and Sarah Sanderson. Of course, my family have always been behind me: the Grovers, Holbrooks, Steeleys, and Wigginses—thank you all for your love and support.

Most of all, I am forever indebted to my husband, Everett Wiggins. Thank you for following me down this trail, helping me climb the mountains, and pulling me out of the mud. Words can never express my love and gratitude, although I will continue to try every day. I love you more, Everett, and I always will.

Generous funding provided by the U.S. National Science Foundation grants 09-43049 and 11-11107 made this research possible. The opinions, findings, conclusions and recommendations presented in this dissertation are my own and do not necessarily reflect the views of the National Science Foundation.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS	iv
LIST OF FIGURES	xiii
LIST OF TABLES	xv
LIST OF APPENDICES	xvi
CHAPTER	
I. Introduction	1
1.1 Citizen Science	1
1.1.1 Challenges and Advantages of Citizen Science	2
1.1.2 Examples of Citizen Science	4
1.2 Problem Statement	7
1.2.1 Research Questions	10
1.3 Concepts	11
1.3.1 Virtuality	12
1.3.2 Technologies	15
1.3.3 Organizing	18
1.3.4 Participation	21
1.3.5 Scientific Outcomes	25
1.4 Significance	26
1.4.1 Audiences	27
1.5 Overview	28
1.5.1 Literature Review	28
1.5.2 Methods	29
1.5.3 Case Studies	29
1.5.4 Theoretical Framework	31
1.5.5 Cross-Case Analysis	31

1.5.6	Conclusions	31
II.	Literature Review	33
2.1	Literature Review	33
2.1.1	Scientific Collaboration	34
2.1.2	Online Communities	41
2.1.3	Public Participation in Scientific Research	47
2.1.4	Citizen Science As Scientific Collaboration	56
2.1.5	Literature Review Summary	59
2.2	Conceptual Framework	61
2.2.1	Inputs	64
2.2.2	Processes	66
2.2.3	Emergent States	67
2.2.4	Outputs	68
2.2.5	Research Questions Revisited	69
2.2.6	Conceptual Framework Summary	70
2.3	Summary	71
III.	Methods	73
3.1	Research Design	73
3.2	Case Selection	76
3.2.1	Theoretical Sampling	77
3.2.2	Selected Cases	82
3.2.3	Comparison of Selected Cases to Broader Population	85
3.3	Case Study Data Collection and Analysis	88
3.3.1	Research Process	89
3.3.2	Data	91
3.3.3	Analysis	96
3.4	Research Quality	101
3.4.1	Validation Strategies	102
3.4.2	Researcher Positionality	105
3.4.3	Going Native	107
3.5	Summary	109
IV.	Mountain Watch	110

4.1	Conservation in the Clouds	110
4.1.1	Project Description	112
4.1.2	Organizational Context	114
4.1.3	Data Collection for Mountain Watch	117
4.2	How Mountain Watch Works	121
4.2.1	Data Contribution Tasks	121
4.2.2	Meta-Contribution Tasks	127
4.2.3	Mountain Watch Technologies	127
4.3	Organizing in Context	134
4.3.1	Organizational Embeddedness	135
4.3.2	Leadership in the Organizational Field	141
4.4	Climate Science for Conservation	144
4.4.1	Science and Participation: Rigor and Revisions	144
4.4.2	Multiple Layers of Data Collection	155
4.5	Volunteers on Vacation	159
4.5.1	Hikers	160
4.5.2	Constant Turnover in Participation	164
4.5.3	Organizing Communication Saturation	166
4.6	Answering the Research Questions	168
V. Great Sunflower Project		172
5.1	Know Bees or No Bees	172
5.1.1	Project Description	173
5.1.2	Data Collected for the Great Sunflower Project	177
5.2	How the Great Sunflower Project Works	180
5.2.1	GSP Technologies	180
5.2.2	Contribution Tasks	181
5.2.3	Meta-Contribution Tasks	187
5.3	Organizing Growth	188
5.3.1	Growing Partnerships	188
5.3.2	Seed Money and Sustainability	191
5.3.3	Size Matters in Organizing	195
5.3.4	The High Cost of Low-Cost Technologies for Participation	197
5.4	Crowdsourcing Conventional Science	200
5.4.1	Evolution of a Scientific Protocol for Public Participation	200
5.4.2	Asking the Right Questions	204
5.4.3	Scientific Outcomes	205

5.5	Participation Experiences for Ordinary People	207
5.5.1	Gardeners	208
5.5.2	Demographics and Data Entry	211
5.5.3	Waiting for the Lemon Queen	213
5.5.4	Organizing Participation: Retention Over Recruitment . .	216
5.5.5	Encouraging Understanding and Intervention	218
5.6	Answering the Research Questions	221
VI.	eBird	224
6.1	Birding for Science	224
6.1.1	Project Description	224
6.1.2	Organizational Context	229
6.1.3	Data Collection for eBird	230
6.2	How eBird Works	235
6.2.1	eBird Functionality	235
6.2.2	Data Contribution Tasks	242
6.2.3	Meta-Contribution Tasks	243
6.3	Organization and Organizing	245
6.3.1	Project Leadership	245
6.3.2	Project Sustainability	249
6.4	Institutional Influences	251
6.4.1	Institutions and Partnerships	251
6.4.2	Broader Institutional Impacts	255
6.5	Designing Technologies for Participatory Science	257
6.5.1	Aligning Scientific and Personal Interests	258
6.5.2	Impacts on Technology Design	260
6.5.3	Impacts on Science Processes	262
6.5.4	Impacts on Scientific Outcomes	264
6.6	Designing Technologies for a Community of Practice	267
6.6.1	Birders	268
6.6.2	Communication is Critical for Organizing	273
6.6.3	Community-Based Participation Recruitment Strategies . .	275
6.7	Participation Incentives and Transformative Experiences	279
6.7.1	Incentivizing Participation	279
6.7.2	Making Data Entry Worthwhile	282
6.7.3	Changing Birder Behavior	285
6.7.4	Individual Development Through Participation	287

6.8	Answering the Research Questions	290
VII.	Theoretical Framework	294
7.1	Theoretical Framework	294
7.2	Complementary Theories	297
7.2.1	Deliberate Design Model from Ecology	298
7.2.2	Strategic Process Model from Social Movements Theory	300
7.3	Project Inputs	305
7.3.1	Scientific Interests	306
7.3.2	Community	306
7.3.3	Resources	307
7.3.4	Institutions	308
7.3.5	Mission	311
7.3.6	Technologies	312
7.4	Individual Inputs	314
7.4.1	Biography	314
7.4.2	Personal Interests	315
7.4.3	Skills	316
7.4.4	Networks	316
7.5	Processes	317
7.5.1	Science	318
7.5.2	Design	319
7.5.3	Organizing	320
7.5.4	Participation	321
7.6	States	322
7.6.1	Sustainability	323
7.6.2	Commitment	324
7.6.3	Satisfaction	325
7.7	Outputs	326
7.7.1	Contributions	327
7.7.2	Individual Development	328
7.7.3	Scientific Knowledge	328
7.7.4	Broader Impacts	330
7.8	Theoretical Framework Reflected in the Cases	331
7.9	Summary	335
VIII.	Cross-Case Analysis	337

8.1	Comparative Analysis of Case Studies	337
8.2	Scientific Work Versus Lifestyle Activities	338
	8.2.1 Revision Cycles	340
	8.2.2 Participation Process Evolution	341
	8.2.3 Design of Self-Reinforcing Experiences	345
8.3	Communities and Practices	347
	8.3.1 Community Integration Versus Community Creation	347
	8.3.2 Aligning Protocols and Community Practices	349
	8.3.3 Community Structure and Volunteer Management	352
8.4	Place and Participation	355
	8.4.1 Technologies and Places	356
	8.4.2 Place Attachment	359
	8.4.3 Ground Truthing	363
8.5	ICT Constraints and Affordances	365
	8.5.1 Affordances: Scale and Communication	366
	8.5.2 Constraints: Usability and Development Costs	367
8.6	Organizing	370
	8.6.1 Organizer-Contributor Ratio	371
	8.6.2 Institutions and Sustainability	372
8.7	Implications	375
	8.7.1 Implications of Case Sampling	375
	8.7.2 Implications for Scientific Outcomes	376
	8.7.3 Implications for Scientific Collaboration	379
8.8	Summary	381
IX. Conclusions		385
9.1	Limitations	385
9.2	How do virtuality and technologies alter organizing in citizen science?	388
9.3	How do virtuality and technologies shape participation in citizen science?	390
9.4	How do organizing and participation influence scientific outcomes in citizen science?	392
9.5	Emergent Findings	394
9.6	Contributions	396
9.7	Summary	397

APPENDICES	399
BIBLIOGRAPHY	422
CURRICULUM VITAE	440

LIST OF FIGURES

Figure

2.1	The overlap of research topics relevant to citizen science.	34
3.1	Variations on paper-based birding checklists.	81
3.2	Research timeline for data collection, analysis, theory revisions, and writing.	90
4.1	AMC hut facilities.	116
4.2	The Mountain Watch visibility monitoring view guide.	123
4.3	Phenology monitoring participation kits for northern forest flowers (left) and alpine flowers (right).	125
4.4	Mountain Watch data sheet drop box at Pinkham Notch Visitor Center, located at an information desk.	126
4.5	Mountain Watch data sheets with plot descriptions and map (top), a location-specific form (left), and generic form for alpine locations (right).	129
4.6	The Mountain Watch online participation forms.	130
4.7	The Mountain Watch online interactive maps at several geographic resolution levels.	132
4.8	Automatic data collection instruments.	134
4.9	Commonly confused alpine pincushion plants.	153
5.1	The Lemon Queen sunflower, <i>Helianthus annuus</i>	174
5.2	The GSP interface for describing a garden.	183
5.3	The GSP paper data sheet for 2011.	185
5.4	The GSP online data entry interface for bee observation samples.	186
5.5	Bee-themed products sold to raise operating funds for the GSP.	194
6.1	Diagram of the accumulation of records in the eBird database over time.	227
6.2	A life list of birds reported on eBird.	236
6.3	An eBird bar graph report of birds seen at a specific location.	238
6.4	A “Top 100” report showing the rankings of eBirders in New York by number of checklists submitted in 2011.	240
6.5	The eBird TrailTracker’s interface for assigning observations to specific locations.	241
7.1	A theoretical framework of citizen science as a type of virtual organization.	296

7.2	The Deliberate Design Model from Shirk et al.	299
7.3	Reproduction of the Strategic Process Model from Ganz (2000).	302
7.4	Mountain Watch concept relationship diagram.	332
7.5	The Great Sunflower Project concept relationship diagram.	333
7.6	eBird concept relationship diagram.	333

LIST OF TABLES

Table

2.1	Factors affecting stages of research collaboration, from Sonnenwald (2007).	36
2.2	Names for different forms of public participation in scientific research. . . .	48
2.3	Stages of the scientific inquiry process that define PPSR models. X = public included in step; (X) = public sometimes included in step.	55
2.4	Research question concepts and related theoretical concepts.	69
3.1	Application of theoretical sampling criteria to selected cases.	82
3.2	Summary of average survey respondent project characteristics and cases. .	86
3.3	Validation criteria and strategies.	103
3.4	Evidence for conformity to evaluation principles for interpretive field research.	104
4.1	Interviewees for Mountain Watch case study	118
6.1	Interviewees for eBird case study	231
7.1	Summarized concepts from the theoretical framework.	295
7.2	Correspondence of concepts from the Deliberate Design Model.	299
7.3	Correspondence of concepts from the Strategic Process Model.	303
7.4	Summary of project inputs from the theoretical framework.	305
7.5	Summary of individual inputs from the theoretical framework.	314
7.6	Summary of processes from the theoretical framework.	318
7.7	Summary of states from the theoretical framework.	323
7.8	Summary of outputs from the theoretical framework.	326
8.1	Association of thematic topics with theoretical framework concepts.	339
8.2	Association of project design approaches with theoretical framework concepts.	339
8.3	Association of communities and practices with theoretical framework concepts.	347
8.4	Association of participation and place with theoretical framework concepts.	355
8.5	Association of ICT with theoretical framework concepts.	365
8.6	Association of organizing with theoretical framework concepts.	370
8.7	Association of thematic topics with research question concepts.	381

LIST OF APPENDICES

Appendix

A.	Sample Interview Protocol	400
B.	Survey Methods	401
C.	Citizen Science Project Profile	403
D.	Citizen Science Project Survey	408
E.	Sample Rich Process Model	421

CHAPTER I

Introduction

This chapter introduces the phenomenon of study, citizen science, and discusses the research problem that the study addresses. The nature of the problem is elaborated in the general and specific research questions, followed by definition and discussion of the primary concepts related to the questions. The significance of the research is then addressed, including identification of audiences for whom the study's contributions may be of interest. The chapter concludes with an overview of the remaining chapters.

1.1 Citizen Science

Citizen science projects involve the public with scientists in collaborative research (Cooper et al., 2007). Many such projects can be viewed as virtual organizations with geographically dispersed resources and members who work toward common goals via cyberinfrastructure. Public participation in scientific research can take a variety of forms. Diverse volunteer populations can contribute to scientific research through a variety of activities, from primary school students engaging in structured classroom projects, to families volunteering together in “bioblast” one-day organism census events, to geographically-distributed individuals monitoring wildlife populations over time. The dominant form of citizen science projects, found in the environmental sciences, focuses on monitoring ecosystems and wildlife populations; volunteers form a human sensor network for distributed data collection (Cohn, 2008; Bonney

& LaBranche, 2004). By contrast, in projects like NASA's Clickworkers (Kanefsky, Barlow, & Gulick, 2001), volunteers provide data analysis service, applying basic human perception to computationally difficult image recognition tasks. Citizen science projects "hold out the possibility of scaling up the processes of scientific research so that they are truly global in scale and scope" (Bowker, 2005, p. 125).

The value of volunteer contributions to scientific research is nontrivial. Schmeller et al. (2009) surveyed 395 biodiversity monitoring schemes led by 227 organizations in 28 European countries, and found that out of over 46,000 individuals contributing more than 148,000 person-days to biodiversity monitoring, just over 13% were professionals. The total cost of these projects was €4 million and the total value was conservatively estimated at €13 million (2006 wages), demonstrating the importance of volunteer engagement for reducing the cost of monitoring species for population studies over large spatial and temporal scales.

1.1.1 Challenges and Advantages of Citizen Science

Despite the increasing popularity of citizen science, there are few reference points to guide project design and implementation. Scientists are consistently and rightfully concerned about ensuring the quality of research outputs, and if public contributions are accepted there are no guarantees of expertise sufficient to establish quality. The simple solution is to permit contributions from only those with sufficient expertise or credentials; given the academic standards for scientific research quality, this represents the traditional scientific research model. Incorporating contributions from a wider population into scientific knowledge production requires additional mechanisms to ensure quality, evident in the details of study designs that use these approaches.

Although the concern over contribution quality is the primary issue raised by such critics as Roman (2009), this is only a symptom of the underlying problem: researchers have not yet established how to consistently select effective tools and mechanisms for contribution and

coordination for the scientific outcomes that they want to achieve. In part, the knowledge gap is because the potential options are not well known. There is currently little guidance to be found on how to select technologies or design participation tasks to ensure the best possible outcomes for both the research and the participants under a variety of circumstances. This problem is exacerbated in the translation of existing practices to technology-mediated citizen science, which requires further consideration of technologies and design choices regardless of whether the project activities are focused on data collection or processing.

The practice of citizen science is often virtual because it frequently involves a combination of spatial, temporal, and physical discontinuities. Increasingly, virtuality means that the interactions of individual contributors are mediated by information and communication technologies (ICT), as opposed to simply being distributed across time and space in the majority of activities, which was the focus of earlier research on virtuality and more representative of citizen science practices prior to the year 2000. Related research on scientific infrastructure, and more recently cyberinfrastructure, underscores the importance of understanding how organizational, task, and technology design requirements interact to influence outcomes (Lee, Dourish, & Mark, 2006; Star & Ruhleder, 1994).

Technology-supported citizen science certainly benefits from economies of scale, but does not always rely upon large numbers of participants for successful outcomes. It is a solution for particular types of scientific research goals, performing best in situations where employing paid professionals is not feasible, usually due to the need to cover large spatial or temporal scales. Citizen science also produces remarkably good results for visual analysis and data processing tasks such as identifying craters on Mars, classifying the shapes of galaxies, or transcription of historic records, and for problem-solving tasks like protein folding. Both visual analysis and problem-solving tasks are computationally difficult, but are simpler and even entertaining for humans. While citizen science is not likely to replace traditional sci-

entific research, it is already finding its place as a complement to professionally-conducted scientific research in fields like conservation biology (Dufour & Crisfield, 2008).

Citizen science is a type of organizational and work design is not new to science, with the Audubon Christmas Bird Count founded at the turn of the Twentieth Century, but cyberinfrastructure and ubiquitous computing now make broad public participation in scientific work a realistic research strategy for an increasing variety of studies. Citizen science is related to long-standing programs employing volunteer monitoring for natural resource management (Ballard et al., 2005; Cooper et al., 2007; Firehock & West, 1995), and is often employed as a form of education and outreach to promote public understanding of science (Osborn, Pearse, & Roe, 2002; Brossard, Lewenstein, & Bonney, 2005; Krasny & Bonney, 2005; Bauer, Petkova, & Boyadjieva, 2000; Spiro, 2004). Citizen science projects are now primarily focused on scientific research with increasing frequency (Bonney & LaBranche, 2004; Baretto, Fastovsky, & Sheehan, 2003; McCaffrey, 2005). The current form of citizen science, which has evolved over the past two decades, places more emphasis on scientifically sound practices and measurable goals for public education than similar historical efforts (Bonney et al., 2009). Ample evidence has shown that under the right circumstances, citizen science can work on a massive scale and is capable of producing high quality data as well as unexpected insights and innovations (Bonney & LaBranche, 2004; Trumbull et al., 2000).

1.1.2 Examples of Citizen Science

A host of volunteer monitoring and citizen science projects are now entirely ICT-mediated, providing access to and for a much wider pool of potential contributors than ever before. The ubiquity of networked technologies enabling large-scale participation creates new opportunities for methodological innovations across a number of scientific fields. Without the affordances of ICT, many of the citizen science projects emerging today would not exist due to the simple economics of developing, implementing, and supporting a research project

that incorporates public participation and is capable of generating scientific knowledge. Two examples are briefly described here to provide grounding for the subsequent discussion of the citizen science phenomenon. The Great Sunflower Project is an example of a citizen science project in which volunteers collect and contribute observational data, and GalaxyZoo represents a type of participation in which volunteers contribute to data processing.

The Great Sunflower Project

The Great Sunflower Project (GSP) was created by a single scientist to study pollinator service. Participating volunteers report data on the activity of bees in their gardens, following a specific protocol for observation and reporting. The project name derives from the Lemon Queen sunflower, carefully selected to support the scientific goals of the project. Participation requires volunteers to grow sunflower plants to maturity in order to report on the activity of bees. Volunteers' contributions are coordinated through a simple website running on an open source content management system, through which participants across North America describe their gardens and report observations. The website also features discussion forums, announcements and news from the scientist, and additional educational content about the importance of bees to our food supply, summarized in the omnipresent slogan, "Bees: responsible for every third bite of food."

What is remarkable about the GSP is its overwhelming success in attracting potential contributors, particularly given the project's meager resources. A few weeks after the scientist sent fifteen emails to Master Gardeners¹ to recruit participation, 15,000 people had registered to participate, well in excess of expectations. The volunteer base has continued to grow, swelling to 77,000 in the summer of 2009 and over 100,000 in the summer of 2011. While the majority of the website registrations do not convert into data-contributing volunteers, the number of contributors is adequate to provide a much larger set of observation

¹The title "Master Gardener" denotes an individual who has completed a program of intensive horticultural training, typically provided by university cooperative extension programs and repaid through outreach-focused volunteerism.

data than could feasibly be generated by professional researchers, and the geographic scope covers the continental United States and Canada. The project has been so successful in attracting volunteers that maintaining project sustainability required changes to the original participation protocol and new fundraising efforts outside of the usual academic funding sources.

GalaxyZoo

GalaxyZoo is a citizen science project organized by an inter-institutional team of professional astronomers. Volunteers apply superior human perceptual capacities to computationally difficult image recognition tasks, providing an important service in data analysis. The classification tasks are performed through a web portal that presents images of galaxies and asks volunteers to make judgments about specific characteristics of the galaxies, with questions such as whether the galaxy has a bulge or a bar in its center, how rounded it is, and perhaps best of all, “is there anything odd?”² The website also includes a blog authored by the astronomers and forums for discussion among participants, providing multiple venues for engagement. In its first instantiation, GalaxyZoo volunteers classified 750,000 galaxies in record time, and the data have been re-incorporated into virtual astronomy observatory tools used by both the public and researchers. In its second version, the GalaxyZoo 2 project elicited far more complex classification judgments from volunteers, implemented based on the high quality of the results from the simpler initial classification. After three years, GalaxyZoo had classified over 56 million galaxies, and had a growing contributor base of over a quarter of a million volunteers.

Beyond simply providing image processing services for science, GalaxyZoo participants have made new discoveries, such as Hanny’s Voorwerp, an astronomical object of unknown nature (*voorwerp* means “object” in Dutch), remarkable for its unusual blue color and for

²http://www.galaxyzoo.org/how_to_take_part

emitting more energy than any object previously observed in the universe. Time on the Hubble Telescope was granted to examine this new astronomical body, which was discovered in 2007 by Hanny van Arkel, a Dutch elementary school teacher. Hanny's Voorwerp demonstrates how profoundly volunteer contributions to scientific research can influence the course of knowledge creation.

In addition to innovation, GalaxyZoo volunteers deliver quality; their reliability is as good or better than that of professional astronomers. The project's leaders ensure quality by having each image evaluated by multiple volunteers, with algorithmic flagging of low-consensus items for professional review. This mechanism for quality control has been used in several citizen science that featured data processing tasks, and can also be applied in some data collection projects (Sullivan et al., 2009). Even without such sophisticated tools and quality assurance strategies, researchers have found that elementary school children can provide scientifically valid data for species identification, with seventh-graders reporting counts of crab species at 95% accuracy and third-graders correctly identifying animals 80% of the time, an acceptable reliability rate for most ecological studies (Cohn, 2008; Delaney et al., 2007).

GalaxyZoo and the Great Sunflower Project are examples of large scale citizen science projects which have found ways to address some of the practical challenges inherent in citizen science. In addition to these issues, the phenomenon also presents theoretically interesting problems, discussed in the following problem statement.

1.2 Problem Statement

Citizen science projects supported by information and communication technologies can yield massive virtual collaborations based on voluntary contributions by diverse participants. The increasing scale of these projects, some of which involve tens of thousands of members

of the public in distributed data collection and analysis, is accompanied by the need for research into the effects of virtuality and technologies on project processes and scientific outcomes. As noted by Silvertown (2009), modern citizen science differs from its historical forms primarily in the access for, and subsequent scale of, public participation. ICT are credited as one of the main drivers of the recent explosion of citizen science activity. To take advantage of this emerging opportunity, scientists need a more thorough understanding of how research design, implementation, and management decisions affect scientific outcomes in citizen science.

In particular, designing information systems to support technology-enabled citizen science requires understanding the effects of organizing and participation processes on the scientific outcomes of citizen science projects. ICT-enabled citizen science projects are similar in some respects to peer production phenomena such as free/libre open source software development (FLOSS) or Wikipedia, but have scientific goals that pose particular constraints on task design. For example, assuring the reliability of data collection is critical to establishing the value of a scientific project, but not a matter that can necessarily be left to the “wisdom of crowds.”

Including volunteers in scientific research projects also results in very different distributed organizational structures than those of scientific laboratories studied to date, raising new challenges. The nature of these collaborative projects is meaningfully different from prior forms of scientific collaboration, as it more closely resembles cooperation than collaboration. For example, creators of scientific laboratories may tacitly assume that participants have comparable and high levels of skill and will contribute relatively equally. This is rarely the case for citizen science volunteers, who may have widely varying levels of skill or knowledge, and contribute at levels differing by orders of magnitude. These projects sometimes bear a greater resemblance to cyberinfrastructure projects than scientific laboratories due to the

larger scale of participation and increased complexity of organizing (Lawrence, Finholt, & Kim, 2007). Combined, these factors raise concerns for designing systems to support citizen science.

Just as there is wide variability in the content and focus of the projects, there is also great diversity in the types of technologies currently implemented to support citizen science. These range from simple open source content management systems to more sophisticated custom software platforms, GPS, and smartphone applications. The use of ICT to support citizen science has already yielded significant impacts on scale and scope of participation and research. Although the sophistication of these technologies is rapidly increasing, most citizen science is still supported by relatively simple, low-cost technology solutions, but there is little guidance to help projects choose and implement appropriate ICT to support citizen science projects' research goals.

Designing and implementing technologies to support cooperative work requires understanding the setting and the nature of the task (Bannon & Schmidt, 1989). The goal of this study is to better understand the processes that these technologies must support and the settings in which participation is carried out. This research motivation suggests focusing on the role of technologies in processes of organizing and participation in citizen science.

In complex settings where the specifics of the context and events are expected to have an important influence on outcomes, such as in citizen science, focusing on processes provides a means for abstraction that permits meaningful comparison across cases (Yin, 1984). Studying processes can illuminate the link between individual and organizational (project) level phenomena:

Viewing a process as the way organizations accomplish desired goals and transform inputs into outputs makes the link to organizational outcomes. Viewing processes as ordered collections of activities makes the link to individual work, since individual actors perform these activities. (Crowston, 2000, p. 38)

This study is focused on developing process theory rather than testing variance theory, in order to better understand how differences in virtuality and technologies influence organizing and participation processes.

1.2.1 Research Questions

Three key observations relevant to the general research questions are drawn from the literature discussed in Chapter II. First, contextual factors require additional consideration in research decision-making when tasks are carried out by unsupervised, distributed contributors with widely variable expertise and skills. Moving citizen science into a technology-mediated format can address some of the logistical constraints encountered by place-based projects, but may do so at hidden costs, and cannot fully eliminate the complexities of designing research to be conducted by distributed, heterogeneous volunteer contributors.

Second, traditional long-term volunteer monitoring practices are being adapted for participation in technology-supported citizen science projects with limited consideration of the consequences of the change in context from face-to-face to technology-mediated participation. The extent of the adaptation for ICT-supported citizen science project deployment often appears to be little more than the direct conversion of traditional paper data forms and training presentations into digital forms and files. Without an established project and existing volunteer base, simply translating these materials into digital versions may not be adequate. As well, digital contribution environments provide a number of affordances that may simplify and improve the mechanisms used to ensure the quality of the research, and although individual project reports often address quality assurance and quality control processes and results, the efficacy of these strategies has not yet been systematically evaluated.

Third, many of the new citizen science projects emerging today would have been unlikely to occur without enabling ICT. There are still successful examples of large-scale projects that do not rely on individual volunteers' use of ICT, typically following a specific tiered

structure of participation in which implementation of a particular protocol is coordinated locally, with data reported back to an umbrella organization or partnership by local project leaders. Most new citizen science projects, however, are eliminating local coordination in favor of centralized coordination via the Internet. While it is technically possible for distributed volunteers to classify galaxies without the use of networked information systems, for example, the material and coordination costs make it practically inconceivable to undertake massive-scale galaxy classification without ICT. Technologies to support citizen science must accommodate the full range of virtuality, from intensely physical participation in some place-based projects to fully ICT-mediated participation in place-independent projects.

These three observations led to the more specific research questions for this study:

RQ₁: How do virtuality and technologies alter organizing in citizen science?

RQ₂: How do virtuality and technologies shape participation in citizen science?

RQ₃: How do organizing and participation influence scientific outcomes in citizen science?

The goal of investigating these research questions is to develop a theoretical framework of citizen science that can inform practice and provide a conceptual foundation for future research. Virtuality and technologies are two aspects of context that are particularly complex, as these characteristics can enable the organizing and participation processes while simultaneously impairing the ability to directly observe these same processes as they unfold. Differences in virtuality, the technologies used to mediate virtuality, and their influences on the interactions between project processes together play a complex role in citizen science, as the following discussion of these key concepts further elaborates.

1.3 Concepts

The primary concepts in the research questions are discussed and defined in this section. The inputs of virtuality and technologies are examined as primary elements of the context

of citizen science. Next, processes for organizing and participation are considered, followed by a discussion of the scientific outcomes that these processes influence.

1.3.1 Virtuality

The social structure of citizen science does not necessarily match definitions of virtual employees, groups, teams, organizations, or communities (e.g., Watson-Manheim, Chudoba, & Crowston, 2002). Instead, these projects can bring together individuals with discontinuous organizational affiliations and memberships, spatial and temporal locations, and work practices. These individuals may still be united by a continuous collective identity and goals. The following discussion conceptualizes virtuality in citizen science as a combination of discontinuities, continuities, materiality and place.

Discontinuities and Continuities

Virtuality is a complex concept and inconsistently conceptualized in the literature, as noted by Watson-Manheim et al. (2002), who focus instead on discontinuities, or “gaps or a lack of coherence in aspects of work, such as work setting, task and relations with other workers or managers” (p. 193). They discuss two dimensions of virtual work that are useful for characterizing virtual organizing and participation: the nature of the virtual work environment, and the aspects of the work that are discontinuous. The discontinuities of work include physical and temporal locations; membership, affiliation and organizational relationships; and cultural aspects. Building on this work, Chudoba et al. (2003) characterize the virtual work environment according to six types of discontinuities: geography, temporality, culture, work practices, organization, and technologies.

These two lists of discontinuities draw on overlapping sets of concepts, from which three seem most useful for better understanding virtuality in citizen science: cultural, temporal, and spatial. In the context of citizen science, the primary cultural discontinuity is likely to be

the role-based divide between professional researchers and laypersons, which is reproduced in work practices and organizational affiliations (aside from the shared affiliation to the common project.) Scientists and volunteers would also be expected to perceive the work practices differently, as work practices are a part of the culture of science that may not be familiar to many volunteers. While many other aspects of cultural diversity may converge in citizen science projects, the scientist/citizen discontinuity is likely to dominate in culture, organizational affiliations, and work practices.

Chudoba et al. (2003) specify temporal discontinuities as specific to collaboration across time zones, although they also mention differing perceptions of time, which seem likely to occur along the same divisions as work practices. Time zones as the primary operationalization of temporality assumes the need to organize around synchronous work, however, which is not the usual case for citizen science. The actual temporal structures of participation are widely variant but are rarely defined by discontinuities across time zones. For the most part, the task structure is asynchronous by design, as the typical citizen science protocol represents a mode of participation that is specifically engineered to reduce interdependencies and frequently requires long-term and repeated task execution.

The nature of the discontinuities of physical and spatial locations can be more complex. Spatial discontinuities reflect the differences of location in which participants do their work, and are one of the great strengths of citizen science. Many projects are conducted via citizen science methods specifically to take advantage of the ability to collect data from geographically dispersed observers. Even in projects where a place is defined by a geographic boundary within which the project is conducted (e.g., within a particular National Park), there may be individual sites for each participating volunteer, leading to a number of separate locations. A diverse array of configurations of spatial discontinuity occur in citizen science, which impose varying requirements on project and technology design.

As Watson-Manheim et al. (2002) note, considering the continuities in a virtual work setting may help reflect on the discontinuities. For each side of the cultural discontinuity between social worlds of scientists and volunteers, the continuities are nearly universal: each side shares goals, values, ICT, and work practices among its own members. Across the cultural divide, however, these continuities cannot be assumed. While shared goals provide continuity across roles, the work practices and scientific values are likely to differ. The technologies supporting the work may also differ, with scientists and project leaders likely to have access to more and different technologies than those used by volunteers. With so many discontinuities and so few continuities between individual participants at a project level, the few continuities that do exist must draw on strong social and narrative infrastructures.

A shared goal or mission is likely as important in citizen science as in most other contexts of voluntary work. At a high level, the shared goals of citizen science are those of science, which are often communicated in an idealized form when presented to the public. Scientific knowledge production is a compelling goal for many individuals who may consider scientific research a public good. Likewise, participation may be perceived as a contribution to improving natural resource management decision-making, for example, linking the explicit scientific research task to a common value related to conservation.

Place and Space

The role of physical place and space in virtuality is often overlooked or assumed to be mediated by technologies, and provides a promising variable for identifying broad types of citizen science participation. Place and space represent a dimension of virtuality that is expected to have significant influence on organizing, project resources, and ICT use. Nearly all observational activities (with a few exceptions) have this quality, as they are structured to collect reports of something in the location in which participants are physically present when participating.

While space denotes a physical location, place indicates meaning attached to a location, and may be more strongly influenced by social than physical dimensions (Hidalgo & Hernandez, 2001; Harrison & Dourish, 1996). Research into the relationships between people and settings, or place attachment, has identified that individuals who are attached to a place are more likely to contribute time and resources to support it, and groups often tap into this emotional attachment to mobilize for protection of special places (Moore & Koontz, 2003). This suggests that place-based citizen science participation is likely to reflect the influence of place attachment.

The composition of place and space, which can encompass a multiplicity of physical and cultural locations, is a fundamental consideration for the design of citizen science activities and technologies. At the extremes, projects that have no place-based element (e.g., GalaxyZoo) exhibit a very different design for the participation processes and supporting technologies than projects that are located within specific places, such as long-term species monitoring at National Parks. Other projects with some elements of place (e.g., The Great Sunflower Project) fall somewhere in between, and may need to support different aspects of participation. The variety in virtuality of place and space in citizen science has interesting implications for citizen science study design and participation. While virtuality is conceptualized in terms of discontinuities of place and space, it is closely linked to the use of technologies that are instrumental in supporting large-scale participation.

1.3.2 Technologies

The Greek root of the word technology is *techne*, meaning neither material tools nor applied skills, but referring instead to knowledge. Technology-as-practice arises from this most basic type of knowledge; modern machines and tools are created as means to ends, and are derived from the subsequent technology (Rojcewicz, 2006). Heidegger identified the dichotomy by differentiating technology as human activity from technology as a contrivance:

We ask the question concerning technology when we ask what it is. Everyone knows the two statements that answer our question. One says: Technology is a means to an end. The other says: Technology is a human activity. The two definitions of technology belong together. For to posit ends and procure and utilize the means to them is a human activity. The manufacture and utilization of equipment, tools, and machines, the manufactured and used things themselves, and the needs and ends that they serve, all belong to what technology is. The whole complex of these contrivances is technology. Technology itself is a contrivance, or, In Latin, an *instrumentum*. (Heidegger, 1977)

Interestingly, these distinctions between definitions of technologies parallel the distinctions between three types of “information” from Buckland (1991): information-as-knowledge, information-as-process, and information-as-thing.

Applying modern usage of the term, from a conceptual standpoint, technologies are any tool or routine that may be used in project processes. A participation protocol, for example, is an important technology for most citizen science projects that are organized around collection of observational data. A paper data sheet is also an important technology that is widely used, even in projects that require online data submission via digital technologies. The term technology is used interchangeably with ICT here to refer to a wide variety of information, computing, and communication technologies that are employed in support of citizen science. In later analysis, however, the distinction between ICT and other types of technologies is specified whenever relevant to the discussion.

What these technologies promise for citizen science is affordable scalability through reduction in coordination and production costs. Silvertown (2009) identified the availability of technologies for collecting and disseminating information from the public as a significant factor supporting the recent explosive growth in citizen science. In particular, involving volunteers in distributed data processing has really only become feasible since broadband Internet access became widely available in developed countries. Developing the cyberinfrastructure to support this type of project requires both adequate initial funding to develop the

tools for participation and fairly large numbers of active participants for an acceptable return on investment. Projects with the resources to develop these custom information systems have had spectacular results, leading to efforts to develop more general, reusable technology platforms to support a larger variety of citizen science projects. The eBird project is a good example of the economies of scale that can be achieved, reporting over 10 million observations collected in 2008, with a cost per observation of approximately three cents, a price that continues to drop as the number of contributors increases (Sullivan et al., 2009).

For long-term citizen science projects, the role of technologies appears to be transformational primarily in the coordinative and communicative dimensions of practice, with wider access and ease of organizing for engaging a larger number of contributors in an online environment. Often the data collection and reporting activities that involve recording data on a paper form are not very different from those involving reporting data via an online form. In fact, many projects combine the use of both paper and digital data forms because volunteers record observations on paper data sheets while in the field, with data entry for online reporting as a direct replacement for (or addition to) faxing or mailing the original documents (Wiggins et al., 2011).

More recently, smartphone technologies enable on-the-spot data submission, offering a promising way to reduce the effort and sources of error involved in volunteer participation (Graham, Henderson, & Schloss, 2011). New platforms combining web portals with mobile applications, such as What's Invasive! and Project Noah, reduce the amount of work volunteers expend on filling and submitting forms. At the same time, these tools can also improve data quality and verification options, reduce lags in reporting with electronic data submission occurring at the time of observation, capture precise location information automatically, and permit inclusion of photographs for expert validation. These variations in the role of technologies in citizen science also blur the line between "online" and "offline" participation,

at least for volunteers who are comfortable with these technologies.

The foregoing discussion focused primarily on uses of technologies to support the task-related processes of participation. Technologies also support citizen science projects with tools for social interaction among participants. Common social computing technologies, such as blogs, discussion forums and social media tools, can permit dialog between participants and may provide an essential function in long-term sustainability for some citizen science projects. Such venues for social participation also provide support for learning and membership processes, the development of collective identity, and the cultivation of a community (or network) of practice (Wenger, 1999). These social processes of participation are mutually constituted with the technologies that enable and constrain them, and the development of these sociotechnical arrangements is the focus of the following section on organizing.

1.3.3 Organizing

Organizing is a simple notion that is operationally complex. Intuitively, organizing is the process of *creating order* (Dictionary.com:2012, 2012a), whether the objects of organizing are artifacts, actions, people, or a combination of these. Organizing also refers to creating a whole from coordinated or interdependent parts, another way of creating order. Order is created when we structure and systematize according to rules. Organizing is a basic human activity: we organize without even being aware that we are organizing.

The intuitive conceptualization is problematic, however, as it promotes an oversimplified view of organizing. The time scale of organizing can vary substantially, sometimes requiring a few moments and sometimes generations. In everyday life, we often organize multiple types of entities according to multiple conflicting rules in both one-time and routine processes. Organizing may be done individually to create order for individual needs, or it may be a social activity imbued with increased complexity from establishing shared meaning.

These social processes of organizing are one focus of the current work. The most common

perspectives in organizational theory focus on organizations as the outcomes of organizing; for example, a classic definition identifies an organization as a set of deliberately created, stable social relations with the explicit intention of continuously accomplishing some specific goals or purposes (Stinchcombe, 1965). From this viewpoint, organizing is conceptualized as the process of generating an organization to meet a particular goal, leading to the notion of organizational design. Organization-focused conceptualizations suffer a failure in terminology that narrows interpretation: effective organizing leads to “organization,” but we may label it in different ways depending upon the objects of organizing, e.g., classification, standards, organization, etc.

A number of prominent organizational theorists conceptualize organizations as information processing systems (e.g., March, Simon, & Guetzkow, 1958; Cyert & March, 1963; Weick, 1969; Galbraith, 1974; Tushman & Nadler, 1978; Daft & Weick, 1984; Stinchcombe, 1990), for which the implicit or explicit goal is reduction of uncertainty through the acquisition and use of information. The expectation that information processing occurs in the service of decision-making reinforces the idea that organizing comprises goal-oriented activity, which is appropriate to the context of citizen science. Taking an information processing perspective provides numerous links to the literature on job and task design, which is relevant to understanding the way that organizing is related to participation.

Adopting a social process view of organizing, Weick (1969) proposes that organizing constitutes organization:

Assume that there are processes which create, maintain, and dissolve social collectivities, that these processes constitute the work of organizing, and that the ways in which these processes are continuously executed *are* the organization. The same processes operate through a variety of media. ... Their appearance may change, but their workings do not. (Weick, 1969, p.1)

The social process perspective increases the range of social structures that may be consid-

ered an organization by focusing on the process of organizing rather than the state of a social structure, e.g., by including articulation work and coordination. This view may be more appropriate for citizen science projects, as they are not usually organizations in the institutional sense. The term “organization” is therefore used throughout this document in reference to formalized social institutions such as businesses and agencies rather than other types of products of organizing.

Several theoretical clarifications on organizing from Weick (1969, p.36–42) are also helpful. First, Weick proposes that organizing processes are continually ongoing because the conditions for operation are continually changing. Social processes never unfold in exactly the same way twice, indicating the centrality of organizing to social activity and emphasizing the recursiveness of these processes. This dynamic perspective recognizes organizing as recursive, characterized by processes of self-production (autopoiesis) that allow a system to adapt to fluctuations in its external environment through changes to its internal organization (Maturana & Varela, 1980; Luhmann, 1995; Mingers, 1995). Second, Weick specifies that control processes in organizations are accomplished by relationships rather than people, which is relevant to citizen science because of the expected influence of cultural discontinuities between project leaders and volunteers. Finally, the aforementioned assumption that organizing is directed toward reducing uncertainty in the informational environment provides a broad conceptualization of organizing applicable to a novel context. More specifically, centralized control and role divisions between staff or scientists and volunteers are specific mechanisms for reducing uncertainty. These mechanisms seem to be a natural feature of citizen science due to the cultural significance of these roles and the expectations of scientific research. The conceptualization of organizing as uncertainty reduction ties the focus on organizing processes in citizen science to the processes of participation through which members of the public are engaged in scientific research, discussed next.

1.3.4 Participation

Participation, simply defined, is “the act of taking part or sharing in something” (Dictionary.com:2010, 2010). By definition, participation links the individual to the collective: the literal Latin translation, “to take part,” indicates a social activity or state in which a person contributes to or is involved in an event, project, or other shared interest. Participation is a central concept in understanding the phenomenon of citizen science, as it is the means through which individuals become involved in collective action. Although the concept of participation in a general sense refers to the activities of all participants in a collective effort, the participation processes of primary interest for this study are those of the volunteers. This document uses the term participation in a similar sense to the way the term “work” is used in the related literature.

Like virtuality, participation is a multi-faceted concept. In citizen science, participation refers to the ideal version of the tasks that are designed and structured to support the project’s research and educational goals, usually documented in a *protocol*, and also refers to the tasks as they are interpreted and actually carried out by volunteers. Participation includes activities undertaken by volunteers to support task-oriented participation and non-task activities such as social interaction. Both of these types of participation contribute to project outputs and are influenced by the technologies supporting project participation, as will be discussed in the remainder of this section.

Task Participation

Task-oriented participation leading to contributions to the scientific goals of a citizen science project takes two forms, ideal and actual. The ideal form of volunteer participation is typically documented in a protocol that provides instructions on how to make contributions to the project. These protocols are variable in detail and formality, and as a genre of

communication between project leaders and volunteers, are usually identifiable by name or inclusion of an enumerated list of steps for participation.

The most common task structures of participation in citizen science feature reduced interdependence between contribution activities. The aggregation of the products of individual-level participation balances out the low task interdependence at the individual level with pooled interdependence at the project level, in which each individual makes a discrete contribution to the whole. The ways that the context of the larger research process is communicated to volunteers may influence individual-level participation and by extension the quality and quantity of the individual contributions. Rettberg (2005) characterizes project-level awareness among contributors to collective narratives as conscious, contributory, and unwitting participation. Citizen science participation falls into the conscious and contributory modes of participation, in which individuals are knowingly contributing to the project, but with varying levels of understanding about the way their contribution fits into the project at the collective level. Promoting such bigger-picture comprehension is a common informal science education goal. The ability of individuals to understand the role of their personal contributions to the aggregated research product is a form of collective attribution that can be specifically supported with tools that help reinforce the value of individual volunteer contributions.

Task-oriented participation is also influenced by the substitutability of individual contributions. In some projects, only one solution or decision is needed (e.g., classification of an image of a galaxy) and redundant contributions provide quality control, while other projects need each unique contribution because they are not necessarily substitutable or redundant (e.g., observations of pollinator service of sunflowers at a continental scale.) Participation in citizen science where redundancy provides quality control bears a stronger resemblance to the innovation-oriented practices of “crowdsourcing” further discussed in Chapter II, in

contrast to projects where geographic distribution of volunteers allows new variations on the well-established practice of volunteer monitoring. Using ICT to support volunteer monitoring does not make it crowdsourcing *per se*, although volunteer monitoring projects can be designed in a way that follows the crowdsourcing model of participation. Instead, citizen science is a phenomenon that is better represented as involving multiple potentially overlapping models of task design and participation.

Social Participation

Taking part in a citizen science project can mean engaging in more than just data collection or processing. Social interactions can be an important aspect of participation, providing a way to motivate ongoing involvement through community development, healthy competition, and knowledge sharing. Social participation can take a number of forms, such as posting to forum discussions or sending messages to email listservs, recruiting friends to participate, or finding ways to engage family in the project activities.

The nature of social participation in citizen science projects is likely related to the extent of virtuality and use of technologies. At a minimum, in some cases certain modes of participation cannot serve as venues for social interaction because the project does not use social technologies to support its activities. It is also possible that technological infrastructure to support social activity among project participants may be implemented independently of the project, either preceding it or emerging in its wake.

The types of social participation that are possible under different conditions of virtuality are not only influenced by technologies, but also by the nature of task. Examining images of galaxies, for example, may be less likely to be undertaken as a social activity than watching sunflowers to count visiting bees, perhaps due to differences in the ease with which the activity can be made a part of existing social routines. Watching for migrating raptors (hawks and eagles) in a colocated, synchronous setting is structured so that the experience almost

certainly involves social interaction during participation, as do local or regional projects that hold group training sessions for otherwise geographically distributed participants. These types of non-task participation can support the development of a collective identity and a sense of belonging (Wenger, 1999), and are therefore expected to be important contributors to the development of a sustainable project. The study examined different forms of task and non-task participation to better understand the role of social interaction in supporting project outputs and sustainability.

Meta-Contribution

A third type of participation is meta-contribution, a form of participation that supports the participation of other contributors (Crowston & Fagnot, 2008). Meta-contribution can be either task-oriented or social in nature, and typically requires more expertise, experience, and tolerance for uncertainty with respect to less structured task types which may require more independent judgment than is involved in the core contribution tasks. Meta-contributors can also represent an added level of hierarchy in role-based social structures, serving to some extent as intermediaries between organizers and volunteers. For example, volunteers who perform quality assurance review of data submitted by other volunteers are providing task-based meta-contribution. Alternately, individuals who answer the questions of other volunteers in forums or provide informal mentorship are engaging in social meta-contribution.

Meta-contribution is seen in numerous other contributory contexts, such as the Wikipedia (Bryant, Forte, & Bruckman, 2005). Creating formalized meta-contribution roles is a common practice in many traditional voluntary work contexts. For example, in train-the-trainer models of organizing, when the trainers who interact with learners are also volunteers, they perform a meta-contribution task. Meta-contribution is of particular interest in the context of citizen science because it has potential to substantially increase the scalability of participation. Although organizers may have limited time to devote to projects, enlisting

the assistance of a wider meta-contributor base can help to support the efforts of a larger number of volunteers than might be effectively managed otherwise.

1.3.5 Scientific Outcomes

Participation and organizing can take a variety of forms that influence the task-related products of a citizen science project and therefore the scientific outcomes. As previously mentioned, expertise is one part of the task design formula, and is consistently the variable mentioned in the concern over quality, demonstrating the practical relevance of identifying effective ways to design appropriate participation processes for diverse participants. The goal of creating a scientifically-valid study that is simultaneously a participation-worthy activity can require design tradeoffs to accommodate these additional constraints. Project outcomes must be evaluated by comparison to the project's own goals, as other contextual differences between projects (e.g., data sharing, funding sources, domain of research, etc.) will likely moderate any quantitative measures of output such as scholarly publications or public data sets. Gathering histories of project development and outcomes over time, however, could surface common course corrections made during project life cycles.

A pilot stage of research design development and adjustment is common across most scientific endeavors that must establish new research processes, but in citizen science projects it is complicated by the uncertainties of incorporating volunteer participation. The study design and piloting process may represent a critical phase for the project's long-term sustainability and scientific production. This has particularly important implications for citizen science projects in research areas that may be constrained to adjusting the protocol once per year, for example, limiting the ability to adapt the study design if early outputs indicate a need for revision.

1.4 Significance

This research contributes to the literature on citizen science, crowdsourcing, and scientific collaboration. It also advances the discussion of virtuality into a new context. The literature on virtuality demonstrates an ongoing interest in the academic research community, and virtuality in group work continues to pose challenges to collaboration success that influence progress in the sciences, evident in continued funding by the National Science Foundation for theory-based research into virtual teams and organizations.

There is also increasing academic attention to large-scale contribution systems and processes of participation in virtual communities. Most prior studies have examined self-organizing systems and peer production, typically focusing on phenomena like open source software or Wikipedia. Citizen science, although similar with respect to the goal of generating collective products through a distributed mode of participation, also displays some interesting points of divergence from peer production as a result of the constraints imposed by the scientific and educational goals of the projects, which will be discussed later. The phenomenon therefore presents an opportunity for research addressing the challenges of virtuality in distributed collaboration contexts that are not characterized by self-organization, providing a meaningful contrast to the prior literature.

The growth and diversity of citizen science projects indicates an opportunity for new research focused on the relationship between organizing, participation, and scientific knowledge production, as well as a need to translate the findings for application to practice. As yet, there has been little published research on the phenomenon of citizen science for several reasons. The first reason is that despite the long history of public participation in science, citizen science is a relatively recent term for a specific form of public participation in science which has been in use for less than two decades. Second, many citizen science projects do

not fit into the hypothesis-testing model of scientific research, and therefore may not be written about in the formal literature (Silvertown, 2009). The shortage of social research on technology-enabled public participation in scientific knowledge production suggests an opportunity for novel contribution to both the literature and practice.

Directly addressing Bannon and Schmidt's (1989, p.369) call for research in computer-supported cooperative work (CSCW) that addresses the "need to develop a theoretical framework that will help us understand the complex interactions between the technical subsystem, the work organization, and the requirements of the task environment" is one goal of this research. The focus on processes in citizen science as the object of conceptual development also provides a theoretical foundation for next-generation technologies and cyberinfrastructures to support this form of scientific collaboration. Further, developing a better understanding of organizing and participation processes in citizen science can benefit practical decision-making and research policy development.

1.4.1 Audiences

The results of this study have two audiences, academic and practitioner. While this research is primarily oriented toward an academic audience, it also seeks to answer questions relevant to the practitioner community. The interests of these two audiences overlap but are not identical, although a significant proportion of the practitioner audience are also trained research professionals.

Academic researchers in information systems, CSCW, and organization studies all stand to gain from the findings in the research areas of virtual organizations, social computing, and social informatics (Sawyer & Eschenfelder, 2002). This research contributes an empirically-grounded theoretical framework that highlights concepts and relationships that may generalize beyond the boundaries of the phenomenon of study, permitting comparisons across different forms of distributed collaboration.

Practitioner audiences are expected to benefit from findings that can support decision-making to improve the design and outcomes of citizen science projects. In particular, examining how different aspects of virtuality impact participation can inform the design of activities and technologies for ICT-supported forms of participation. Investigating organizing and participation processes and outcomes provides an opportunity to identify best practices and effective uses of technologies to support both task-related and social aspects of participation.

1.5 Overview

This overview of the remainder of the study briefly summarizes the following chapters.

1.5.1 Literature Review

The literature review begins with literature relevant to the context of the phenomenon of interest, followed by description of an initial conceptual framework. The basis for the research is grounded in literature on public participation in science, scientific collaboration, and online communities. The review provides an overview of scientific collaboration, including discussion of collaboratories and cyberinfrastructure that support distributed scientific work. Studies of online communities find evidence of high quality collaboration outcomes from diverse contributions from volunteers, with self-organizing peer production structures that share some features with citizen science. Moving to the literature focused on citizen science, the review summarizes descriptions of several forms of public participation in science, focusing on typologies of participation in citizen science more specifically. Many of these typologies center on the nature of the collaboration between scientists and volunteers, highlighting the design of participation opportunities.

The review is followed by an introduction to an initial conceptual framework for research on citizen science. The framework conceptualizes the phenomenon at both individual and

project levels, including concepts and relationships expected to be relevant to citizen science. This initial conceptual framework was based on literature from small groups theory and additional literature from a variety of fields, attuned to the contextual factors expected to have the most influence on the phenomenon, and was used to guide data collection and analysis.

1.5.2 Methods

The research design is a comparative case study, with the conceptual framework introduced in Chapter II providing the focus for data collection and analysis. Selected concepts from the framework provided the primary concepts for further investigation, and theoretical sampling was employed to select three cases for in-depth study. Field research methods were used to collect several types of data to test and further develop the theoretical framework. As data were collected, analysis began with interview transcript coding and description of each case for within-case analysis. Comparisons were drawn throughout the data collection and analysis process, with the within-case analysis completed before cross-case analysis was undertaken in earnest. Combined with the specified research questions, the iterative and concurrent data collection and analysis strategy employed both inductive and deductive approaches. The quality of the research is strengthened by several elements of the research design.

1.5.3 Case Studies

Each of the three chapters describing the case study projects provides background on the project, discusses how it operates, analyzes the emergent themes and relationships evident in the project, and relates these observations to the concepts from the research questions.

Mountain Watch

Mountain Watch is a citizen science project operated by the Appalachian Mountain Club (AMC) with two sub-projects focusing on air quality and plant life cycles (phenology.) Based primarily at the AMC facilities in New Hampshire's White Mountains, phenology monitoring is the primary focus of Mountain Watch, which is also designed for data collection in other forests and mountain ranges in the Northeastern U.S. Mountain Watch represents a maturing, long-term, place-based citizen science project that has demonstrated rigorous scientific approaches to protocol refinement and has leveraged organizational resources to expand outreach to a constantly changing participant base.

The Great Sunflower Project

The Great Sunflower Project engages participants across North America to answer research questions that are important to understanding and protecting pollinator populations. It was founded in 2008 by Dr. Gretchen LeBuhn, an academic researcher at San Francisco State University. For her, organizing a citizen science project promised a larger and more geographically diverse data set to support her research, as well as an opportunity for public outreach and education. The project's initial research questions focused on bee visitation rates across urban, suburban, and rural habitats. The Great Sunflower Project is a young, underfunded, and technologically disadvantaged citizen science project that has shown remarkable potential and resilience despite substantial challenges.

eBird

eBird is a popular citizen science project operated by the Cornell Lab of Ornithology, a nonprofit organization focused on bird conservation and research and an international leader in developing and promoting citizen science practices. eBird allows users to keep birding observation records online and users can submit data by completing online checklists of birds

seen and heard while birding. The data have been used for policy development, conservation and land management decision-making, countless tools and reports for birders, and scientific research across several disciplines. eBird is widely considered one of the most successful citizen projects in existence, and represents a mature, well supported, and technologically sophisticated project that has engaged volunteers internationally on a large scale.

1.5.4 Theoretical Framework

This chapter presents the theoretical framework that was iteratively developed throughout the study. It served as a lens for focusing the research and identifies several practical considerations for citizen science projects and can help direct future research. The theoretical framework is presented with a systematic review of each concept, with examples drawn from the cases.

1.5.5 Cross-Case Analysis

The emergent findings from the case studies include five thematic topics that relate to theoretical concepts from both the framework and the research questions.

1. Citizen science project design approaches that favor science versus hobbies for participation design.
2. Project design and organizing implications of engaging communities of practice.
3. Relationships between physical environment, technologies, participant experiences, and data quality.
4. Information technology tradeoffs: helpful for scale and communication, challenging for usability and resources.
5. Resources and sustainability relate to institutions and scale of participation.

1.5.6 Conclusions

This chapter begins by discussing the limitations of the research. It then reviews how the foregoing chapters answered the research questions and highlights additional relationships

between concepts from the theoretical framework. It also suggests opportunities for future research and outlines the contributions of the study.

CHAPTER II

Literature Review

Citizen science is increasingly popular as a means for broader audience to engage in scientific work. While information and communication technologies (ICT) are not strictly necessary to accomplish many goals of citizen science projects (one could imagine submitting observations by US mail, and indeed, this mode of contribution is still in practice), the economies of scale from using ICT are such that many of the projects emerging today would not have happened without the enabling information technologies. It is therefore important to consider how technologies and virtuality influence processes and outcomes in citizen science.

This chapter reviews the literature in several areas relevant to citizen science, developing a foundation for investigating citizen science as a type of virtual organization. The literature provides grounding in the phenomenon of citizen science which informed the development of the conceptual framework. The following section introduces the conceptual framework used to guide the research process and the evolution of the empirically-based theoretical model presented in Chapter VIII.

2.1 Literature Review

Literature selected for review has particular relevance to the context of citizen science as a type of virtual organization. It is an unusual blend of adjacent topics: distributed scientific

collaboration, online communities, and citizen science. These areas of research overlap with the others, as shown in Figure 2.1. However, they have not yet come together as a triad, leaving a gap in the literature that will become increasingly significant with more growth of scientific work taking place through online communities. This review discusses each of these topics in turn and then summarizes the literature to provide a contextual foundation for the development of a conceptual framework to guide further study of the phenomenon.

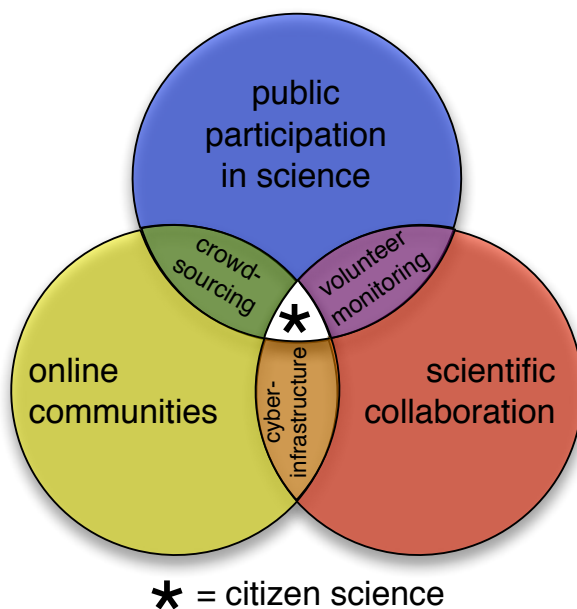


Figure 2.1: The overlap of research topics relevant to citizen science.

2.1.1 Scientific Collaboration

The research literature investigating scientific collaboration is enormous, sprawling across a variety of disciplines. Its relevance here is to establish the broader context within which citizen science operates through discussion of scientific collaboration processes and governance structures, collaboratories, and cyberinfrastructure. More attention has been given in the research literature to collaborations among scientists than collaborations involving the public. This is not surprising, given that the dominant model of science in the last century

has favored an ideal of formal education as the qualification for producing scientific research. However, as public participation becomes an increasingly valuable component of scientific research, a reorientation of the perspectives from these studies becomes necessary to understand citizen science as an evolving form of distributed scientific collaboration. This section briefly reviews the topic, with an overview of scientific collaboration processes more generally, distributed collaboration more specifically, and finally cyberinfrastructure and eScience as the phenomena linking scientific collaboration with online communities.

Scientific Collaboration Processes

Sonnenwald (2007) provides an excellent overview of the research on scientific collaboration, which spans many fields and appeals to a variety of research interests. The review defines scientific collaboration as social interaction among scientists in the interest of furthering a common goal, noting that the tasks involved in scientific collaboration have a high degree of uncertainty that are complicated by functional and strategic dependencies between researchers. Most literature on scientific collaboration is focused on discontinuities in culture, place, and organization, resulting in rich streams of research on disciplinary diversity (e.g., Qin, Lancaster, & Allen, 1997; Cummings & Cross, 2003), geographic distribution (e.g., Finholt, 2002), and organizational and community relationships in scientific collaboration (e.g., participatory action research).

Sonnenwald (2007) organizes a more detailed discussion of the extant research according to the stages of the scientific collaboration lifecycle, which include foundation, formulation, sustainment, and conclusion. The factors affecting each of these broad stages of research collaboration are considered in greater depth throughout the article, and are included in Table 2.1, as these factors represent important concepts to consider in developing a conceptual framework to study scientific collaboration. Notably, these stages of scientific collaboration can be interpreted as a process model, in which research progresses through the four stages

Stage	Components	Factors
Foundation	Antecedents to collaboration	Scientific; political; socioeconomic; resource accessibility; social networks & personal
Formulation	Initiation and planning of collaborative research	Research vision, goals & tasks; leadership & organizational structure; ICT; IP & legal
Sustainment	Work must be sustained over time to achieve goals	Emergent challenges; learning; communication
Conclusion	Collaborative results emerge	Definitions of success; dissemination of results

Table 2.1: Factors affecting stages of research collaboration, from Sonnenwald (2007).

in overlapping sequence. What the categorical structure of the classification fails to represent is that scientific collaboration is often a series of ongoing relationships punctuated by multiple collaborative projects, leading to a looping structure in which conclusions lead to new formulations and foundations.

The governance structure of scientific collaboration has taken several forms. Chompalov, Genuth, and Shrum (2002) characterize the organizational styles of scientific collaboration as bureaucratic, semi-bureaucratic, and participatory, noting that the latter category was applicable in only one field, particle physics, where the factors for the foundation stage of collaboration were markedly different from other disciplines. Upon examining the relationship between governance forms and research practices, “the major connection that emerges is between the structure of leadership and the character of interdependence—greater interdependence leads to decentralization of leadership and less formalization” (Chompalov et al., 2002, p.752). This finding provides an interesting contrast to self-organizing online communities where work is also considered highly interdependent, but is differently structured. In open source software development, for example, resource-specific dependencies are all but eliminated so that lower interdependence permits decentralization of leadership and less formalization.

Leaving aside concerns of governance structure in collaboration, Olson and Olson (2000) identify specific challenges that contribute to success or failure of research projects, drawn

from the research on distributed collaboration. These challenges are more broadly applicable, and include common ground, work coupling, collaboration readiness, and collaboration technology readiness. More generally, these challenges represent the social, coordinative, cultural, and technological barriers to collaboration. Studies of collaboration often highlight social and cultural issues, for example, the challenges that arise in research engaging partners from developed and developing countries who have differential access to resources (e.g., Cohen, 2000). Other studies focus on the challenges encountered in promoting data sharing, which plays a key role in scientific research collaboration but is strongly influenced by social and cultural norms and practices. Birnholtz and Bietz (2003) examined how data contribute to scientific fact and scientific community in three research domains. They found that data contributed to the scientific community functions of boundary management and status indication while also providing a vehicle for increased engagement. These functions of data are particularly relevant for consideration in the context of citizen science, as volunteer participation often takes the form of data contribution or reduction. Likewise, volunteers' access to resources for participation may be a concern for some citizen science projects with resource dependencies related to technology-mediated participation.

The methods for studying scientific collaboration are nearly as varied as the aspects of the phenomena that they study. The literature on scientific collaboration includes many bibliometric studies employing analysis of citation and authorship patterns to understand knowledge production in the academic sphere. The focus of these methods on the knowledge artifact as evidence of collaborative behavior is not well suited to understanding collaboration that involves the public in science. For example, the hundreds of thousands of contributors to eBird are not co-authors on the academic papers produced from the shared data, but are still critically important collaborators in the production of scientific knowledge and are recognized in the acknowledgement sections of articles. Instead, other forms of evidence

are required to examine the nature of scientific collaboration when the research engages a broader participant base than scientists alone.

Collaboratories

Combining the words collaboration and laboratory, the term “collaboratory” denotes a virtual research environment in which ICT provides distributed collaborators access to colleagues, instrumentation, shared data, and computational and intellectual resources (Wulf, 1993). The move towards collaboratories for scientific work is predicated in part on the increase in large-scale “big science” research (de Solla Price, 1963), and the idea of distributed intelligence mobilizing scientific effort via ICT (Finholt, 2002). Besides investigating the information technologies that can support complex work, another goal of research on collaboratories is reducing geographic and status barriers to interaction between scientists, as increasingly affordable ICT have become ubiquitous in scientific work. While collaboratories partnering with formal education environments have shown promise (Finholt, 2002), these efforts focus on providing an avenue to participation rather than a means of increasing capacity for scientific knowledge production.

Although the early research on collaboratories focused almost exclusively on collaboration among scientists, The Science of Collaboratories project conducted a landscape survey of IT-enabled research collaboration (Bos et al., 2007). This research demonstrated that virtual environments have promoted contribution to scientific research by a wider variety of participants. The Science of Collaboratories project developed a typology for describing a variety of distributed scientific practices. Among the seven types of collaboratories, two engage the public: Community Data Systems and Open Community Contribution Systems models. The differentiation between these two forms reflects a practice-based distinction between public participation in data collection versus data processing tasks. Together, these models describe the majority of technology-enabled citizen science projects.

Much of the research on collaboratories focuses on the tools used to support distributed collaboration. For example, Farooq et al. (2007) examine the technology design requirements for CiteSeer, a large-scale digital library of scientific literature. As the scale and complexity of a scientific collaboration increases, the literature has shifted from describing these projects as collaboratories to cyberinfrastructure projects, although there is no definitive separation between these forms. Researchers have returned to exploratory and descriptive research designs to begin to understand the situated contexts in which cyberinfrastructures are developed and used.

Cyberinfrastructure

Cyberinfrastructure (CI) for the sciences takes collaboratories to the next level; the term refers to “the coordinated framework of technology and human expertise intended to support scientific discovery, particularly research requiring high-performance computing, large quantities of data, or distance collaboration” (Lawrence et al., 2007, p.1). The related concept of eScience refers more specifically to the practices of distributed collaboration that are reliant on cyberinfrastructure (Hey & Trefethen, 2005). Citizen science projects often rely on a combination of technologies and human participation, can produce scientific cyberinfrastructure in the form of research resources, and are often characterized by spatial discontinuities among participants. As such, citizen science can be considered a type of eScience as well as a form of cyberinfrastructure: it is both a set of practices representative of a particular type of distributed scientific collaboration, and its enabling technologies and products can provide a form of sociotechnical foundation for scientific research characterized by large data sets, distance collaboration, and high-performance computing. Research in the areas of cyberinfrastructure and eScience are currently in their infancies, with a few initial studies focusing on the same challenges as those of collaboratories, particularly coordination, geographic dispersion, and social aspects of sharing in science.

Lawrence et al. (2007) identified several differences between cyberinfrastructure-based projects and collaboratories in the literature, as the magnification of geographic, organizational, and cultural discontinuities of project members created a number of new challenges. Unlike collaboratories, CI projects were found to have decentralized leadership and less flexibility due to tightly coupled work, because the growing scale of participation results in increasing discontinuities and heavier reliance on ICT to moderate their effects. Citizen science also increasingly relies on ICT to overcome discontinuities inherent in massively distributed work, often with the goal of expanding the scale of participation. Although the scant research on cyberinfrastructure to date assumes that scientists are the primary participants whose work must be supported, the situation that is presented in these studies is very similar to the conditions of technology-mediated citizen science due to the considerations of scale and scope.

The shift toward cyberinfrastructure-based organizing has impacts on social aspects of scientific work. Lee et al. (2006) discuss the role of coordination and social practices in developing cyberinfrastructure, while De Roure et al. (2008) focus on the technical aspects of supporting social interaction and sharing among scientists in a scientific virtual research environment, myExperiment. myExperiment is just one of a variety of sophisticated tools available to enable open science, which involves sharing research and analysis products throughout the research cycle, promoting high transparency in scientific research. As David, den Besten, and Schroeder (2006) note, however, the complex social structures and incentives of the institutional arrangements within which science is conducted means that eScience does not necessarily equate to open science. These and other social and institutional factors that form barriers to adoption for collaboration technologies can be compounded by infrastructural barriers as well (Star & Ruhleder, 1994). Social support for eScience in the form of cyberinfrastructure such as myExperiment's social network platform may not nec-

essarily provide a readily-adopted route to open science, but it does offer a clear example of the link between scientific collaboration and online communities more generally, to which the discussion now turns.

2.1.2 Online Communities

Prior research has examined a variety of online communities, which thrive based on voluntary contributions in various forms from members at large. Ellis, Oldridge, and Vasconcelos (2004) review the literature on virtual communities, identifying four key themes: virtual communities of practice, virtual arenas, and virtual communities based on either proximity or common interest, with the latter being well established for scientific communities. These studies offer useful reference points for comparison to better understand the nature of participation in technology-mediated citizen science projects. Online communities research examines the nature of social structure and participation in virtual environments, a perspective missing in the literature on citizen science, which has focused to date on the social structure of participation more generally. Research into online communities that engage in knowledge production often characterizes the work as self-organized peer production, featuring progressive engagement of individuals who are motivated to become members of a community of practice through ongoing contribution. Most online communities in the literature resemble self-organizing peer production models; however, citizen science typically follows a more hierarchical model. This contrast makes the prior literature useful for better understanding the differences and similarities between these types of contribution communities, as the remainder of the section demonstrates.

Self-Organizing in Online Communities

In studies of online communities, researchers have identified a number of instances where self-organization is a primary feature of the social structure. When Markus, Manville, and

Agres (2000) discussed rules and institutions as a mechanism contributing to functional virtual organizations, they specified the adaptability and customizability of self-governance. These features are uncommon in citizen science because protocols and rules for participation are often centrally controlled and non-negotiable. In citizen science, the ability to engage in monitoring and sanctioning of contributors is also usually restricted to those in leadership roles, rather than distributed among members of the community and enacted via social mechanisms as in other online communities. The management of membership generally works differently in most citizen science projects, where professional qualifications may be required for advancement to membership as a core contributor. Of the four mechanisms from Markus et al. (2000), reputation alone serves the same role in citizen science and open source software development, and volunteer management efforts of all stripes take advantage of the motivational aspects of reputation. A notable difference in virtual work environments is the improved ease of counting and algorithmically ranking contributors, enabling immediate feedback to reinforce desirable contribution behaviors.

Building on learning theory, Wasko's theoretical model of knowledge production highlights the role of technologies in translating a face-to-face social structure into a virtual social structure and the adjustments of social mechanisms to fit the computer-mediated context (Wasko, Faraj, & Teigland, 2004). The assumptions of self-organization in the practice-based associations, reflective of those observed by Markus et al. (2000), do not hold for the full range of citizen science projects. While some citizen science projects may resemble these structures, the majority are hierarchically structured. Often the only true self-organization that volunteers engage in is self-selection—the choice to participate or not—and in some cases, how much to participate.

Similarly, Crowston et al. (2007) examine self-organization in open source software development, finding that self-assignment to tasks is a key coordination mechanism used in the

entirely distributed production environment. Self-assignment to task stands in stark contrast to task assignment practices in citizen science, which in any environment typically involves standard protocols for participation. Unlike open source developers, volunteers do not define or select their own tasks except on a limited basis as permitted by protocols. This is a result of the scientific goals of citizen science projects, which nearly always demand uniform data collection and analysis procedures. The protocol-based approach seems most similar to the task assignment mechanism that Crowston et al. (2007) identify as “assign to an unspecified person.” Because this task assignment mechanism gives the same task to all volunteers, it cannot be considered self-organization: someone has to define the task. From the opposite perspective, however, self-selecting to participate in a given project represents a functional form of self-assignment to task in citizen science when we consider that a volunteer has the choice to participate in other projects with different task structures. The fact that the task is uniform for most participants effectively makes all volunteers peers in the production process; however, as the next section will discuss further, it does not mean that they are engaging in peer production as discussed in the prior literature.

Peer Production

In the literature on peer production systems, motivation is used to explain the choice to participate (e.g., Wilkinson, 2008; Viégas, Wattenberg, & McKeon, 2007). Most citizen science participation, however, does not meet the definition of peer production (Benkler, 2002), in which the prototypical model is generally non-hierarchical and self-organizing. Neither characteristic describes the typical citizen science project. Nonetheless, a theoretically-focused analysis of collaborative peer production provides useful dimensions for evaluating differences in contribution (Haythornthwaite, 2009). Examining the task design, social structure, and reward structure in the lightweight peer production model focuses attention on a feature of the work design which is also evident in citizen science projects: pooled interde-

pendence, in which each incremental piece of work contributes to the whole without being contingent on other parts. Pooled interdependence and incremental work appear to be key features of effective task design in citizen science. These dimensions of work practice provide a basis for understanding the integration of individual participation processes into the project level in citizen science knowledge production processes.

Returning to the phenomenon of lightweight peer production discussed by Haythornthwaite (2009), one aspect that is particularly relevant to citizen science projects is the weak tie association among contributors, which is more aptly described as coorientation to a common enterprise than as a true community structure. As a whole, lightweight peer production strategies rely on minimally sized and minimally complex work units that are completed by large numbers of contributors (Haythornthwaite, 2009). This mode of work makes up for the inconsistency of participation and continual turn-over of contributors to maintain sustainability despite dynamic membership (Butler, 2001). These principles of work design are also consistent with the project design requirements for citizen science, which expend effort to support volunteer retention but may also need to acquire greater numbers of contributors to achieve geographic scale in addition to making up for inconsistencies in participation. Attracting sufficient participation to make lightweight peer production effective is a matter of mobilizing a large number of volunteers, which in turn relies upon volunteers' motivation to participate, considered in the following section.

Motivation and Engagement

Citizen science project participation resembles the lightweight peer production model (Haythornthwaite, 2009), which provokes the perennial questions about motivation to participate. While participation in peer production is generally expected to be motivated by self-interest, citizen science projects appear more altruistic on the surface. In practice, this perception seems partially true. Participants in GalaxyZoo reported multiple motivations

which reflect both altruism and self-interest. Raddick et al. (2010) found that the dominant motivations for contribution most strongly emphasized an interest in astronomy and fascination with the vastness of the universe, which seem to be self-interested motivations, but the second most common motivation was a desire to contribute, which reflects a potentially altruistic motivation as well. Raddick et al. (2009)'s discussion of citizen science practices also emphasizes potential social benefits arising from progressive levels of engagement in citizen science.

General models of progressive engagement are echoed elsewhere (e.g., Preece & Shneiderman, 2009; Fischer, 2002). Core-periphery models of voluntary participation, much like those seen in research on traditional work groups (Cummings & Cross, 2003), are a consistent feature across a number of these domains, such as open source (e.g., Crowston et al., 2006). Even simple models of progressive engagement, moving from periphery to core, or novice to Wikipedian as observed by Bryant et al. (2005), demonstrate role-based contextual differences that contributors experience in this transition. Related work develops a theoretical model of the motivational arc of contribution in massive virtual collaborations, with separate models for motivations for initial versus sustained contributions (Crowston & Fagnot, 2008). The usual focus on motivational factors does not seem to adequately credit the role of experience resulting from participation, however, which may transform the perceived benefits of initial participation into experienced benefits for continuing participation.

While these studies consistently find that only a few contributors will advance through the ranks into more engaged roles, there is not often such an advancement structure in place for citizen science projects despite the frequent emphasis on individual development. Volunteers are usually free to do more repetitions of their specified task, but most participation processes do not provide avenues for volunteers to engage in additional steps of the research process. While some projects are structured in a way that invites individual inquiry to de-

velop alongside of the structured processes of participation, most projects are based on tasks that do not offer a clear way for individuals to extend their efforts beyond the confines of the protocol. This is also a common feature of crowdsourcing efforts that restrict public participation to very specific tasks, as discussed next.

Crowdsourcing

Crowdsourcing is an ill-defined but increasingly common term which refers to a set of distributed production models. It is typically used to describe an outsourcing strategy that makes an open call for contributions from a large, undefined network of people (Howe, 2006). Initially introduced as a novel alternative business model, attention has turned more recently to the application of crowdsourcing practices to a variety of problems in other domains, including those that produce social goods and scientific knowledge. Early definitions of crowdsourcing revolved around the role of corporate entities in drawing on the “wisdom of crowds” (Surowiecki, 2004), but more recent popular use of the term has applied it to any form of collective intelligence that draws on large numbers of participants through the Internet. Although the specific definitions and application of the term are as yet contested, the broader practice of crowdsourcing links citizen science with online communities.

Travis (2008) reports on InnoCentive’s use as a platform for supporting crowdsourcing for nonprofits and public goods, but in many scientific contexts, doubts as to the value of crowdsourcing arise, primarily regarding veracity and accuracy of crowdsourced research products (Roman, 2009). However, these concerns often overlook the fact that the design of the crowdsourced task must be appropriate to the scientific goals and heterogeneity of contributors in order to generate scientifically valid outcomes, as in any other scientific project that enlists volunteers.

The discussion now turns from scientific collaboration and online communities to the literature focused on citizen science projects, the context for this study.

2.1.3 Public Participation in Scientific Research

Citizen science is related to a number of phenomena across a variety of areas of research and practice (Bonney et al., 2009), such as volunteer monitoring, community science, living labs, and participatory action research. The boundaries separating these practices and defining the space for citizen science are fuzzy, as scholars in different fields have used the same terminology to refer to different types of participation. This section begins with descriptions of the various labels that have been applied to different practices related to public participation in research. This discussion of models of participation provides a basis for understanding citizen science as a phenomenon and a contextual foundation for the conceptual framework developed later in this chapter.

Definitions of public engagement in science have produced conflicting definitions which are used differently across communities of scholars and domains of practice. Table 2.2 shows the labels for eight types of public engagement in scientific research, along with the research domain where they are used and the key features of each practice. The different terminology also represents different contexts of research practice (e.g., environmental versus behavioral sciences) and varying degrees of public involvement in scientific research. The variations highlighted by these overlapping terms are primarily related to the ways the public participates in the research.

The different labels for these practices are grouped into three categories in Table 2.2, representing general classes of participation based on the type of public engagement in the scientific endeavor. Using terms from Lawrence (2006), these are described as consultative, functional, and collaborative forms of public participation in scientific research (PPSR.) The common features of each of these categories will be discussed in the following sections, which also provide definitions for the terminology and contexts of use.

Type	Label	Research Domain	Key Features
Consultative	Civic science	Science communication	Public participation in decisions about science
	People's science	Political science	Movements for people-centered science
Functional	Citizen science	Environmental sciences	Public participation in scientific research
	Volunteer monitoring	Natural resource management	Long-term monitoring and intervention
Collaborative	Participatory action research	Behavioral science	Community participation for action
	Action science	Behavioral science	Participatory, emphasizes tacit theories-in-use
	Community science	Psychology	Participatory community-centered social science
	Living labs	New product development	Public-private innovation partnerships

Table 2.2: Names for different forms of public participation in scientific research.

Consultative Participation

Although the usage of terminology related to PPSR varies considerably, the term generally refers to a different set of phenomena than those grouped under the headings of Public Engagement in Science (PES) and Public Understanding of Science (PUS). In PES, citizen participation is typically policy-oriented and consultative, while PUS usually has connotations of outreach and education, and is primarily studied by scholars of science communication. The defining characteristic of PPSR is that lay people participate in doing science. In PES and PUS, they react to or engage in decision-making about science policy, or learn about science and scientific outcomes without direct engagement in scientific work.

In particular, *civic science*, a term related to both PES and PUS, refers to civic efforts by professional scientists to communicate with the public about science, as opposed to science that is designed and carried out by the public (Irwin, 1995). Scientists who play the role of *civic scientists* in science communication are distinguished from the lay public by formal education and vocation (Clark & Illman, 2001). An excellent contemporary example of a civic scientist is Neil deGrasse Tyson, an astrophysicist who has had substantial influence

on public perceptions of science and has been named one of the ten most influential people in science for his role in popularizing science (Kruglinski & Long, 2008). Civic science is concerned with increasing public participation in science from a general standpoint and with democratization of scientific knowledge production (Bäckstrand, 2003).

In an extension of the discussion on civic science, Irwin (2001) questions the policy-oriented construction of public understanding of science, in which he refers to participants as *scientific citizens*, whose involvement in *scientific citizenship* is limited to consultative and informative aspects of a deliberative democracy model of participation focused on outreach for science education and engagement in policy-oriented activities. In practice, public engagement in policy has taken on a number of forms, such as public opinion surveys, consensus conferences, and deliberative democracy initiatives. Inviting the public to direct scientific research with the goal of addressing the interests of the people is a largely Western practice, originating with policymakers and researchers. The same concept of engaging the public in directing science to meet the people's needs has taken a very different form in India, where it has sparked social movements.

People's Science Movements (PSM) are a form of scientifically-oriented social movements that vary significantly from one group to another in the degree of public participation in scientific work. In the 1970's, PSM emerged as a public reaction to elitism of science in India. The anti-elitism sentiment led to the emergence of organizations with widely varying scopes and scales, ranging from local groups focused on research responding to particular local needs to mass movements advocating "science for social revolution" (Vaidyanathan, Krishnaji, & Kannan, 1979). Although most PSM were instigated and supported by organizational entities, rather than emerging from the populace, they are unquestionably social movements that seek to involve local citizens in scientific affairs to address their own problems.

As the movements developed, however, the terminology was clarified to specify that "peo-

ple’s science” refers not to the science that the people have, but rather the goal of bringing the scientist’s science to the people. Instead of a focus on science developed from traditional and indigenous knowledge or local participation in the production of scientific knowledge, it focuses on bringing scientific knowledge to the indigenous people and helping science and technology researchers understand the problems of India’s impoverished people (Kumar, 1984).

Consultative participation in science, as represented by the concept of civic science, India’s PSM, and other specific practices, typically constrains the public’s engagement to awareness and policy-related purposes. The active participation of the public in contributing directly to scientific research is a form of functional participation, which is described next.

Functional Participation

The goals of functional PPSR focus on engaging non-scientists in the scientific research process to produce scientific outcomes. Like consultative participation, there is great variety in the practical instantiations of functional participation and the use of terminology to describe these practices.

Irwin (1995) defines *citizen science* as scientific research that is initiated and completed by members of the public. At around the same time, the term was independently coined by Rick Bonney at the Cornell Lab of Ornithology to describe a form of research collaboration involving the public in scientific research to address real-world problems (Bhattacharjee, 2005; Bonney et al., 2009; Cohn, 2008). Bonney’s definition of the term citizen science does not require that the research is initiated by members of the public. Unlike Irwin’s definition, this interpretation has taken root in the larger practitioner community among both researchers and volunteers through usage in practice, where citizen science has become an important approach to addressing a genre of scientific research problems.

The typical citizen science model of PPSR “engages a dispersed network of volunteers

to assist in professional research using methodologies that have been developed by or in collaboration with professional researchers” (Cooper et al., 2007, p. 2), with the explicit expectation that volunteers are involved primarily in data collection, and that the inquiry addresses researchers’ questions rather than questions developed by the volunteers. A related definition for *citizen scientists* defines the role of the public as “volunteers who participate as field assistants in scientific studies” (Cohn, 2008). Notably, these projects are primarily scientist-led, and this form of contributory citizen science is the focus of the current study.

The term *volunteer monitoring* is now essentially synonymous with citizen science, but is historically best known in the applied domain of natural resource management, where it has been practiced in North America for over a century (Firehock & West, 1995). In these projects, lay persons are trained to make scientific observations for long-term monitoring, typically of natural resources. Cooperation between volunteers and researchers is the practice that links scientific collaboration with PPSR in Figure 2.1.

In large-scale volunteer monitoring projects, public participation is usually limited to data collection activities structured by scientists, although in more localized contexts it can serve as a basis for action (Firehock & West, 1995; Fernandez-Gimenez, Ballard, & Sturtevant, 2008). While most strongly associated with watershed research, volunteer monitoring is a term that has been applied more broadly, e.g., to monitoring of invasive species, wildlife population, and weather. Watershed monitoring projects are more likely than other volunteer monitoring efforts to originate as grassroots efforts from citizens organizing scientific inquiry to address a common problem (Danielsen, Burgess, & Balmford, 2005; Savan, Morgan, & Gore, 2003).

These functional forms of PPSR, citizen science and volunteer monitoring, are seen in a diverse range of projects with scientifically-oriented goals. Unlike consultative PPSR, the public is functionally engaged in doing research, although usually for only one or two steps of

the research process. When members of the public are engaged in most or all of the research process, these practices are forms of collaborative PPSR, which will be discussed next.

Collaborative Participation

Collaborative participation is very different from consultative and functional models, as the members of the public (more often referred to as community) are engaged in most or all of the process of scientific inquiry. Collaborative participation positions the power relationship between scientists and citizens according to a radically different model than is associated with traditional science, and typically intends to promote empowerment and direct action. Participatory action research, community science, and living labs are examples of public participation in science that engages the participants as collaborators in the research process. These collaborations typically have goals of social action and technology development. Most collaborative approaches are variations of action research, a reflective and often iterative process of problem-solving in which individuals (scientists) lead a community in research focused on action-oriented outcomes.

Participatory action research (PAR) is a methodology representing a much broader category of critical research. It differs from action research by involving members as subjects and co-researchers (Argyris & Schon, 1989). The goal of PAR is understanding and improving the world through change, and it is distinguished from other research approaches through its focus on enabling social action, its careful attention to power relationships, and its dynamic approach with respect to involving the social group being researched as researchers (Baum, MacDougall, & Smith, 2006). The application areas for PAR are broad, ranging from human-computer interaction to public health, but this approach carries a number of drawbacks for social research because it is difficult, in practical terms, to strictly adhere to the basic tenet of fully collaborative research, in which the community under study is engaged in every step of the research process (Walter, 1998).

Community science is a distinct vein of action research in the field of psychology, which seeks to develop a science of communities in which the communities are collaborators in knowledge production, similar to PAR but with a specific emphasis on integration of research and practice (Wandersman, 2003). The goal of community science is to develop knowledge about human behavior in community contexts, to the benefit of the community members (Tebes, 2005). Contextualization is a core value of community science research (Luke, 2005). Yet another subtle variation on this category of research is known as *action science*, which is a form of action research that emphasizes investigating and documenting the theories-in-use that come from the participants (Argyris & Schon, 1989). Action science takes its cues from practitioner perceptions in local practice-based contexts; the methodology tests theories through interventions that attempt to both test hypotheses and create a desirable change.

A relatively new form of PAR branches into research involving private sector entities in *living labs*, which represent a type of situated experimental approach involving “users/consumers/citizens” in innovation-driven partnerships to develop products and services, most frequently information technologies (Eriksson, Niitamo, & Kulkki, 2005). Living labs research has seen academic applications (e.g., Intille et al., 2005), but is primarily in use in Europe as a form of research and development practice that typically focuses on ubiquitous computing, mobile technologies, and collaborative work-support systems. Under ideal conditions, living labs research involves academics, end users, communities, and companies in the creative process of user-centered design and evaluation in product development through cooperative co-creation and study of technology use in naturalistic settings (Almirall & Wareham, 2008). Private sector involvement by business firms clearly differentiates living labs from other forms of public participation in research, although like civic science, the living labs approach calls for public involvement in funding, organizing, and governance (Niitamo et al., 2006).

In summary, the definitions for practices related to citizen science invoke substantially different concepts of participation, ranging from consultative to collaborative, and operate in a wide variety of contexts. The primary feature by which these practices are typically differentiated is the degree of involvement of the public in research processes. The way that specific practices are characterized by researchers reflects more deeply upon their paradigmatic perspectives, discussed in the following section.

Typologies of Citizen Science Participation

The labels and definitions from the prior section describe a wide range of social practices that demonstrate the diverse roles of the public in scientific research. The discussion now turns to three typologies from the ecological sciences that attempt to describe these practices, focusing primarily on functional and collaborative modes of engagement.

The engagement of public participants is examined in greater detail by focusing on engagement in different steps of scientific research. The level of detail in these analyses differs, as do their final categorizations, but they are largely in alignment. Table 2.3 lists the different steps in scientific inquiry used in the definitions for three classes drawn from a recent report from a comprehensive assessment of participation models in citizen science.

Besides evaluating the stages of scientific inquiry in which the public is involved, Cooper et al. (2007) include additional details of geographic scope and research, education, and management goals, which are contrasted in six models of scientific inquiry. The goal of the typology is to present a framework for integrating individual property owners in monitoring and active conservation efforts in residential areas. The research models represented by the typology demonstrate the interplay of scientist and landowner roles in adaptive management practices, which apply scientifically informed natural resource management strategies in an iterative process of intervention, evaluation, and revision.

Distinguishing *community science* from *citizen science* (in yet another dual usage of termi-

Stage of Inquiry	<i>Contributory</i>	<i>Collaborative</i>	<i>Co-created</i>
Define question			X
Gather information			X
Develop hypotheses			X
Design study		(X)	X
Data collection	X	X	X
Analyze samples		X	X
Analyze data	(X)	X	X
Interpret data		(X)	X
Draw conclusions		(X)	X
Disseminate results	(X)	(X)	X
Discuss results & ask new questions			X

Table 2.3: Stages of the scientific inquiry process that define PPSR models. X = public included in step; (X) = public sometimes included in step.

nology) based on the community control of the inquiry, Wilderman (2007) proposes an alternate typology that includes community consulting, community-defined research, community workers, and community-based participatory research. These categories are congruent with those presented elsewhere with two exceptions. First, Wilderman (2007) differentiated between two forms of community workers models based on whether or not analysis activities are exclusive to scientists. Second, this typology also included a category in which the community is engaged in a consultative capacity, represented as “science for the people,” much like the consultative participation discussed earlier. The contrast of consultative practice against “science by the people” casts the typical scientist-initiated citizen science project development model in a negative light; an alternate perspective might suggest “science with the people” as another potential model to consider. In addition, the analysis discussed differences in efficiency, sustainability, and democracy of projects as distinguishing features of these participation models.

Bonney et al. (2009) provide a comprehensive, educationally-focused technical report that summarizes many of these views. The report discusses contributory, collaborative, and

co-created PPSR projects, shown in Table 2.3, synthesizing prior typologies. The authors examined case study projects with a rubric-based evaluation to make a multi-faceted assessment of outcomes in several key areas of focus. The final framework resembles a simpler variation on the models from Cooper et al. (2007) and Wilderman (2007), but includes more detail with respect to the steps of scientific inquiry in which volunteers may be included, moving the sophistication of the typology up a level despite its apparent simplicity. The role of the public as collaborators in scientific research is clearly important to the practitioner community, as it is the primary theme underlying each typology of citizen science practices. The current study focuses on the contributory model of citizen science, as it dominates current practice.

2.1.4 Citizen Science As Scientific Collaboration

The prior sections have presented a review of prior literature on scientific collaboration and public participation in scientific research. The definition of citizen science used in this study refers to these practices as a form of collaboration. The nature of collaboration in citizen science is, however, notably different from that of traditional scientific research conducted in small groups. It is, in fact, fairly comparable to scientific collaboratories and cyberinfrastructure projects. Research on collaboratories had shown that in addition to reshaping the traditional roles of collaborators in scientific research, “some domains of activity are more naturally inclined toward collaboration (data collection vs. contemplation and idea formation)” (Finholt, 2002, p. 95). In citizen science, carrying out the scientific work of data collection or processing is the primary role of participants, which has led to the question of whether this is truly collaborative science. The position taken in this study is that contributory citizen science is indeed collaborative science, given changing notions of what constitutes scientific collaboration.

The essential question is what specific parts of the scientific work a person must engage

in to be considered a collaborator. To clarify a semantic point, all collaborators in scientific research projects are participants in scientific collaboration. Therefore, the simplistic separation of roles into “organizers” and “participants” in citizen science can be reframed in terms of “investigators” and “assistants” in conventional scientific collaboration. In traditional scientific collaboration, assistants are considered part of the collaborative structure, although credit-giving mechanisms are widely available.

Projects that fit the criteria of collaborative or co-created projects, as discussed in the prior section, are intuitively more comparable to traditional scientific collaboration than the contributory projects that are the focus of this study. Based on typologies of citizen science, the typical view is that when members of the public are involved in a wider range of the steps of the research process, this makes participants collaborators rather than glorified sensors. Indeed, this is the fundamental difference between the categories of public participation in scientific research that are discussed as “functional participation” versus “collaborative participation.” The classification hinges on the question of whether participants are contributing in a functional role versus a more intellectual role characterized by participation in a broader array of scientific tasks.

This is a slippery slope: in traditional laboratory research, for example, graduate students often play a parallel role to that of citizen scientists in contributory projects, with little influence on the intellectual agenda despite making a fundamental contribution to the actual completion of the supporting work. The lab assistants in such a scenario may participate in only one step of the scientific process, or several research assistants may carry out identical tasks that require little or no individual judgment. The only substantive differences between these assistants and most citizen science participants is the putative level of training and the degree of variability in work tasks across individuals; even these differences may not be so great in some situations.

At the project level, however, a more meaningful difference lies in the organizational complexity of the research collaboration. Citizen science typically demonstrates a simple overall structure with a relatively flat hierarchy of roles that directly govern the types of tasks individuals undertake and the interdependency of their work. Despite this simplicity, citizen science projects are typically organized in a bureaucratic style with clear role divisions based on assumed expertise (Chompalov et al., 2002), while cyberinfrastructure projects follow a variety of organizing styles. Cyberinfrastructure projects typically exhibit complex social structures that cannot be adequately described by traditional organizational structures, distributed teams, or individual networks (Lee et al., 2006). In these instances, the “human infrastructure of cyberinfrastructure achieves collective action not by making my relationship to the whole visible but by making it invisible, indeed irrelevant. The human infrastructure does not create a distributed team; it dissolves the very need for one” (Lee et al., 2006, p. 491). Despite the relative difference in overall complexity of the scientific work, citizen science shares these properties with cyberinfrastructure.

As large-scale scientific research requires an ever-increasing number of individuals, the contributions of each individual take on similar properties to citizen science participants with respect to the nature of the work as discrete and constrained to a single step in the research process. Finholt (2002) notes that the increasing scale of scientific collaboration in terms of the numbers of collaborators (as measured by authorship, which relies on progressively less reliable assumptions about collaborative practices) is pushing scientific collaboration further toward a model of “distributed intelligence”. In the distributed intelligence model of scientific collaboration, exemplified by cyberinfrastructure projects, the nature of the work is more like citizen science than the traditional conceptualization of scientific collaboration as occurring in small groups. This suggests that studying citizen science may produce findings with implications for scientific cyberinfrastructure projects.

2.1.5 Literature Review Summary

The context of this study is grounded in literature on scientific collaboration, online communities, and public participation in science. The review described several forms of public participation in science and focused on typologies of participation in citizen science. Many of these typologies focus on the form of collaboration between scientists and volunteers, leading to a brief review of the research on distributed scientific collaboration. The literature is dominated by studies of scientific collaboration between scientists, with little attention to the increasingly well established practice of engaging members of the public in research.

Virtual modes of contribution make it possible for a broader audience to engage in scientific work. An increasing number and variety of citizen science projects are taking advantage of the affordances of information technologies to advance scientific research. The forms of participation usually involve contributing data according to an established protocol or completing structured recognition, analysis, or problem-solving tasks that depend on human competencies. With thoughtful study design, contributory styles of functional participation can generate reliable, valid scientific outcomes.

Citizen science represents a type of distributed scientific collaboration. The stages of scientific collaboration in research suggest that such a project will build upon the broader environmental conditions during the foundation stage to provide the necessary ingredients for the formulation stage, in which a new scientific collaboration is planned and started. In most citizen science projects, these stages are the domain of professional researchers and the volunteers enter the collaboration in the sustainment stage, during which the project's work is carried out over time. In the conclusion stage, during which dissemination and evaluation of success occur, a broader notion of success may be appropriate in citizen science. Until only recently, the majority of research on scientific collaboration focuses on projects in which all contributors to the collaboration are scientists or supporting professionals. In

the Science of Collaboratories project, however, two types of collaboratories describe the majority of technology-enabled citizen science, differentiated by whether they are structured around open contributions consisting of data or of analysis. With the growth of scientific collaboration, emerging scientific cyberinfrastructure projects represent a phenomenon that bears stronger similarity to citizen science than the collaboratories discussed in the literature to date.

Citizen science is clearly a different way of organizing online contribution than has been analyzed in the literature. Unlike most online communities that have been studied, these projects are not self-organizing. They do not generally represent peer production in the same sense as other knowledge production networks because they are nearly always hierarchically structured, with scientists designing the work that volunteers then contribute. The structure of tasks is similar, however, and existing literature on the task structure in peer production can inform practice for citizen science. The nature of community is varied across contexts, but the hierarchical aspects of citizen science projects would tend to create a different sense of community with respect to authority, leadership, and decision-making. On the other hand, the typical core-periphery structure of many online communities is not dissimilar from a structural standpoint. Finally, there are strong similarities with respect to issues of motivation and progressive engagement that bear a striking resemblance to virtual communities or networks of practice, albeit with scientists as overseers of the community's practices. This suggests that the models of motivation and participation from studies of online communities of practice may provide insight into the design of tasks and technologies to support citizen science communities. The discussion now turns to description of a conceptual framework that provided initial focus for this research within the context just outlined.

2.2 Conceptual Framework

This section describes a conceptual framework that integrates the literature reviewed in the previous section with findings from an empirical pilot study to guide the design of the main study (Wiggins, 2010; Wiggins & Crowston, 2010). The pilot study examined the evolution of a new citizen science project that was developed by a regional partnership to study phenology, the natural life cycles of organisms, and provided empirical context for the initial conceptual framework. As is discussed later in Chapter III, the framework was applied theoretical sampling and guiding the initial stages of data analysis, and generated the resulting theoretical framework presented in Chapter VII through iterative cycles of revision.

This framework conceives of citizen science projects as a kind of small group, specifically, a work team. Guzzo and Dickson (1996) defined a work team as “made up of individuals who see themselves and who are seen by others as a social entity, who are interdependent because of the tasks they perform as members of a group, who are embedded in one or more larger social system (e.g., community, or organization) and who perform tasks that affect others (such as customers or coworkers)” (Guzzo & Dickson, 1996, p.308). A team differs from a community of practice because members have a shared output whereas in communities of practice (e.g., the copier repairmen studied by Orr (1996)), members share common practices, but are individually responsible for their own tasks (Wenger, 1999). In communities of practice, members also share specialized vocabulary, contextualized learning, and sociocultural ways of understanding that originate in community practices. Notably, while taking part in a community of practice is typically represented as tasks carried out independently, tasks may also be shared and interdependent, involving multiple community members (Hutchins, 1995).

Members of a citizen science project are likely to share goals and social identity, and per-

form interdependent tasks. Although these tasks are typically designed to reduce reciprocal and sequential interdependencies to reduce coordination costs, conjunctive interdependencies remain, so collective outcomes are strongly affected by pooled interdependence (Thompson, Zald, & Scott, 2003). Even though individual tasks seem independent, the final product is the pooled contributions, and value of each individual contribution is dependent upon the totality of contributions to the pool. This differs from prior literature in organizational settings in that citizen science project members may vary greatly in their degree of identification with and contribution to the project, so these factors should not be taken for granted.

The initial conceptual framework represents an early stage of theory development, and will be revised with the addition of empirical evidence that supports or disconfirms the model. Standard organizational forms can be overly simplistic or otherwise inadequate as a basis for understanding organizing in citizen science. Taking this perspective strengthens the framework by better accommodating the wide variety of ways that projects organize their activities. The project level of group interaction is distinct from those of small work groups and organizations (Grudin, 1994), which has implications for the types of information technologies employed to support group activities. Adopting the project rather than the organization as the unit of analysis does not impose assumptions about organizational arrangements. Project teams and communities of practice can be distinguished by their goal orientation among other features (Wenger, 1999).

The framework draws on work in the small group literature, (e.g., Hackman & Morris, 1978; Marks, Mathieu, & Zaccaro, 2001; McGrath & Hollingshead, 1994), incorporating concepts and relationships from the literature on organizational design, job design, volunteerism and participation in virtual communities, at both individual (i.e., volunteer, staff member) and project levels. The perspective taken in this study is substantially different from prior work on citizen science models, which focused primarily on participation structures with

little regard to organizing or processes. Therefore, the initial framework was not strongly influenced the prior literature on public participation in scientific research at a conceptual level, although some of the contextual aspects of the domain were incorporated through literature on volunteerism in particular.

Synthesizing elements from organizational design, sociology and studies of nonprofit management with small group theory strengthens the understanding of the antecedents of scientific knowledge production through massive virtual collaboration. Given the potential of citizen science to operate similarly to other forms of massive virtual collaboration such as open source software development, the framework was adapted from a model developed from a review of empirical literature on open source software development to extend an earlier input-process-output (IPO) framework (Crowston et al., 2005).

The conceptual framework is organized as an input-mediator-output-input (IMOI) model (Ilgen et al., 2005). An IMOI framework was chosen because it provides a general structure for developing a model of socially-embedded groups over time, and improves on the prior IPO models of work groups by including feedback loops and separating the moderators of emergent states from processes. In this framework, inputs are the starting conditions of a team, which includes member characteristics and project/task characteristics (Hackman & Kaplan, 1974). Mediators represent factors that affect the influence of inputs on outputs and are further divided into two categories: processes and emergent states. Processes represent dynamic interactions among members as they work on their projects, leading to the outputs. Emergent states are concepts that characterize dynamic group properties, which vary based upon context; they describe the group's cognitive, motivational and affective states, rather than activities and processes. Outputs are the task and non-task consequences of the system functioning; although there are conceptual differences between the terms "output" and "outcome", these terms are used here interchangeably. The feedback loop from outputs to

inputs treats outputs as inputs to ongoing processes and emergent states; as a result, not all processes or inputs may be active at any given time, depending on the state of system functioning.

The concepts in this framework simultaneously represent the system at both individual and project levels. The model was also customized to include contextual factors that were expected to have the most salience to this particular phenomenon, citizen science, while retaining some level of comparability to other IMO models. The following sections present a general overview of the inputs, moderators, and outputs that served as the initial focus for this research.

2.2.1 Inputs

Inputs are the starting conditions of a project, including both individual-level characteristics and project-level characteristics. At the individual level, staff and volunteers come to the project with diverse demographics, levels of skill, and motivations for participation that affect their individual contributions to the project. While demographics and skills will vary among volunteers involved in different projects, both practical reports and academic theory suggest a number of common motivators for volunteerism, which may have differential effects on individual experiences and performance (Lawrence, 2006; Pearce, 1993; Cnaan & Cascio, 1999).

Four concepts (purpose, environment, resources, and technologies) combine empirical observations from the pilot study with the original expectations that people, technologies, and project design decisions are important factors influencing the emergence and ongoing operation of a citizen science project. The purpose for a project is also highlighted as a key factor, influencing organizing and research processes as well as individual incentives for participation. The concept of purpose also relates to the goal orientation of contributors and members of communities of practice. Environment acknowledges the importance of the physi-

cal and virtual spaces in which participation occurs. Likewise, resources are included because a variety of material and social assets are required for a successful project. Resources are considered an individual-level input, because theorizing resource and information flows that are not tied to individuals can lead to inappropriate assumptions about the organizational structure of the project. As organizational theorists have observed, “organizations do not have mechanisms separate from individuals to set goals, process information or perceive the environment. People do these things” (Daft & Weick, 1984, p.285).

A design perspective that focuses on the way decisions are made in project planning and operation is implicit in the model as a dimension of the concepts of purpose, community, environment, and technologies that form a foundation for participation in citizen science. Technologies are inclusive of the broader set of processes, tools, and infrastructure supporting citizen science projects. Design and use of technologies are of particular interest given the potential of cyberinfrastructure to support numerous aspects of citizen science. Best practice guides recommend that project partnerships include a scientist and an educator to address the scientific and educational goals of the project, and a technologist to address potentially substantial data management and information systems challenges (Bonney & LaBranche, 2004; Chin & Lansing, 2004). When considering how project design and task design interact with cyberinfrastructure in the context of scientific collaboratories, the entire research process must be examined. For example, concerns for the usability of data reporting forms (and subsequent usefulness of the data) has prompted some emphasis on usable technologies and interfaces for volunteers. Understanding the range of interactions between such diverse end users and technologies that support the participation and scientific research processes is important to creating usable, robust cyberinfrastructure systems for collecting useful independent contributions by distributed volunteers (Luther et al., 2009).

2.2.2 Processes

In an IMOI model, the inputs are conceptualized as influencing the effectiveness of projects through two sets of moderators, processes and emergent states. Processes are the dynamic interactions among group members leading to outputs. In this context, volunteer involvement can vary widely, from primary school students engaging in structured classroom projects to geographically-distributed individuals monitoring wildlife populations over time. Understanding these work practices is the first key to designing technological and social arrangements that support knowledge production in virtual organizations of citizen scientists. Notably, project-level processes are accomplished through individual-level processes such as participation and organizing.

At the project level, the processes of interest include those of scientific research itself. The nature of the research and discipline has an important influence on the kinds of data and analysis required and the mapping of tasks to different actors, e.g., volunteers and professional staff. Similarly, data management processes could have a significant impact on project outcomes, particularly in interorganizational projects that must ensure interoperability and reliability of data created by volunteers.

Communication within a project is reflected in different forms of participation and organizing. Participation encompasses the range of activities contributing to citizen science, including task-based, social, and meta-contributions made by individuals whose work supports the efforts of other contributors. An interesting aspect of this context is the applicability of volunteer management processes often associated with nonprofit management, e.g., recruitment, selection, orientation, training, supervision, evaluation, recognition, and retention of volunteers (Pearce, 1993). Organizing processes establish the design and management of a project at the initiation stage and then continue to reshape interactions on an ongoing basis.

2.2.3 Emergent States

Emergent states are dynamic properties of the group that vary as a function of inputs and processes; past research suggests a number of potentially relevant emergent states. The category of “emergent states” is relabeled as simply “states” in later discussion, as they can only be emergent at an initial stage but the IMOI model emphasizes an ongoing system. These include task-related factors that describe the state of the group in terms of its progress on the scientific task, as well as social factors that describe social states of the group that enable that work (Lee et al., 2006). At the level of the project, research on other kinds of virtual organizations has identified the importance of factors such as trust, cohesion, conflict and morale that affect the sense of group community, and thus its long-term sustainability (Markus et al., 2000).

At the individual level, the evolution of volunteers through different roles in the group, from initial volunteer through sustained contributor, and potentially to more central roles, is relevant to project design. A related concern is volunteers’ level of commitment to the project and how it influences their task performance (Cnaan & Cascio, 1999). Understanding how these factors affect the social and technological barriers and enablers for participation is important for effective cyberinfrastructure and project designs.

Collective identity is a concept employed in a variety of literatures, including social movements theory and learning theory (e.g., Lave & Wenger, 1991; Wenger, 1999; Gotham, 1999). It refers to a state of common identification within a group, an aspect of community membership which is recognized and shaped by both members and non-members (Daft & Weick, 1984). The development of a collective identity at the project level is expected to be an important factor in relation to individual commitment, roles, and motivations. Collective identity may also have an impact on the scientific outcomes of a project as well as non-task outcomes like project sustainability.

2.2.4 Outputs

Outputs represent task and non-task consequences of a functioning group, signaling effectiveness. At the individual level, the task outputs are contributions, often raw or processed data, although other types of contributions are possible. In addition to the individual-level outputs, a citizen science virtual organization will have outputs at the project level, such as the scientific knowledge created from the data. Innovative findings, processes, and tools can also emerge from involving the public in scientific research. For example, a new astronomical body, called Hanny's Voorwerp, was discovered by a Dutch elementary school teacher volunteering with the GalaxyZoo project (Cho & Clery, 2009).

Hackman (1987) also includes non-task outputs in the model of group effectiveness. Satisfaction of individual participants' needs, such as individual learning and personal satisfaction, are measures of effectiveness closely related to the educational goals of many citizen science projects. Hackman also includes the group's continued ability to work together, speaking to the sustainability of project goals and social structure. In other words, virtual organizations and citizen science projects are not effective if they achieve a goal but drive away participants in the process.

An important feature of an IMOI model is that outputs themselves become future inputs to the dynamic processes. Positive personal outcomes can lead to increased motivation for future participation, and individual learning can increase a member's ability to contribute. At the project level, learning may lead to innovation in research approaches, resulting in changes to the task design and group processes. Positive project outputs may lead to increased interest among scientists in engaging the public in research and increased visibility for the project, helping to recruit and retain additional volunteers. At the societal level, the success of a project may affect public participation in and perception of science, create informal learning opportunities, and enable knowledge production at an unprecedented pace

RQ concept	Framework	Dimension
Virtuality	Input: Environment	Spatial Temporal Physical
Technologies	Input: Technologies	Task support Social support
Organizing	Processes: Organizing, Research	Project development Research design Coordination
Participation	Processes: Participation	Idealized Task Social
Scientific Outcomes	Outputs: Knowledge, Innovation	Research products Discovery New/revised approaches

Table 2.4: Research question concepts and related theoretical concepts.

and scale (Trumbull et al., 2000; Cohn, 2008).

2.2.5 Research Questions Revisited

Returning to the research questions posed in Chapter I, which focus on the influence of virtuality and technologies on organizing, participation, and scientific outcomes of citizen science projects, some additional directions for research are suggested by bringing these theoretical elements together. Table 2.4 connects the conceptual framework with more specific dimensions of the concepts from the research questions, as previously discussed in Chapter I. According to the categorization from Shaw and Jarvenpaa (1997), these concepts represent a mix of events and variables that may be more or less predictable in the specifics, but are unpredictable overall.

The specific dimensions of the concepts presented in Table 2.4 provided additional focus for the study. Suggesting links between abstract theoretical concepts and operational dimensions to help direct inquiry, these concepts indicated potential sources of evidence for addressing the research questions.

A more specific example demonstrates the way that Table 2.4 can be translated into data collection and analysis: participation is associated with idealized, task, and social processes.

These aspects of participation provide a starting point for further study of the link between participation and scientific outcomes. For example, instructions for participation (protocols) can be evaluated to identify idealized participation processes and triangulated with interview reports from project leaders. Participant observation can be combined with interviews to better understand both task-oriented and social participation processes. Analysis can begin by identifying and describing these processes in order to compare idealized and actual participation, searching for points of divergence in the ways people with different roles understand project participation and evidence of the influence of these processes on scientific outcomes.

The conceptual framework was used to advance the study of organizing and participation in citizen science by employing the concepts highlighted in this section as an initial source of focus at the beginning of the study. Throughout the process of data collection and analysis, this conceptual framework was used to continue testing and evaluating the theory against empirical findings and plan next steps in the research. The comparative case study discussion in Chapter VIII will present a revised version of this framework that more accurately describes the way that projects unfold in practice.

2.2.6 Conceptual Framework Summary

The framework discussed in this section represents an initial conceptualization of citizen science as a virtual organization. The framework was built from a foundation in small groups theory and elaborated with additional literature from a variety of fields, as well as empirical data from a pilot study. The incorporation of contextually-specific concepts and processes related to the phenomenon, such as project-level processes of scientific research and organizing, are linked to individual-level participation in citizen science by relationships suggested by the literature. While many of the individual concepts discussed in this chapter may be more broadly generalizable to other forms of open collaboration, a few elements focus on the science-oriented nature of the phenomenon, incorporating important aspects of the context

without prescribing a particular type or form of citizen science. The outcomes speak to the goals typically associated with contributory citizen science, and the mediating processes and emergent states link the inputs and outputs that lead to the successful achievement of outcomes.

Subsequent chapters describe how this conceptual framework was applied to guide data collection and analysis. The resulting framework developed in this study represents the primary contribution of the research to both theory and practice, as it is empirically grounded and more specific and comprehensive than prior related models. The current study led to iterative revisions to this framework to develop a theoretical model (discussed in Chapter VIII) that can guide deductive research on citizen science and may have value for application to other contexts of virtual participation.

2.3 Summary

This chapter discussed the literature supporting this study from both contextual and conceptual perspectives. Reviewing the research on scientific collaboration and online communities highlighted ways in which citizen science projects fit into the intersection of these areas of study. Examining the variety of practices that relate to citizen science more broadly provided background to contextualize the conceptual framework. In particular, reviewing the literature related to public participation in scientific research highlighted the diversity of historic and existing practices. It also provided a foundation for the focus of the current study on the contributory model of citizen science in which participants contribute to a science-led project.

Conceptualizing citizen science projects as a type of work group while incorporating concepts from the organizational literature indicated that a project-level focus was a valuable distinction from other levels of aggregation by which the phenomenon could be examined.

Building a conceptual framework from the general inputs-moderators-outputs model began with an existing model for a phenomenon with some similar qualities. The framework was then elaborated and refined by integrating concepts from the literature relevant to citizen science and initial refinements based on a pilot study for the current work. The research methods that were guided by the initial conceptual framework are the focus of the next chapter.

CHAPTER III

Methods

This chapter discusses the methodology for the study, first presenting the research design, followed by case selection criteria and a summary of data collection and analysis procedures. The chapter concludes with a discussion of strategies used to ensure research quality.

3.1 Research Design

Technology-supported citizen science is a complex sociotechnical phenomenon. While there are some existing typologies of citizen science projects (e.g., Bonney et al., 2009; Wilderman, 2007; Cooper et al., 2007), there has been little social research conducted to explain how science works when the public is a key participant, much less when the work is ICT-mediated. The goal of the current study is to describe the phenomenon of citizen science, refine the conceptual framework, and develop empirically-grounded theory to describe the conditions, processes, and products of citizen science projects.

The focus of this research is the project organizers, rather than project participants. As a result, data collection focused primarily on gathering information about the management of these projects, and interviews focused only on project organizers and related staff. There were several reasons for the choice to focus on organizers.

Most prior empirical studies of citizen science focus on participants to the exclusion of organizers. In many ways, the participants are an obvious focus for researchers interested

in individual-level theories. For this reason, many avenues for theoretical development with respect to citizen science have already been explored in other contexts. Additionally, many citizen science projects also conduct their own internal evaluations by eliciting feedback from participants, so pursuing a participant-focused direction would not necessarily lead to substantive new insights.

Drawing attention instead to the way that projects are created and managed can complement the prior work and provide a different perspective on the phenomenon. An important goal for this study is to support further research into the phenomenon, and on a practical level, to produce findings that can support improvements to project practices. Therefore, the focus of the study is on the complex project-level systems that enable and contextualize individual participation. While the decision to focus on organizer perspectives sacrifices some breadth, it permits greater depth with respect to the comparisons that can be drawn.

The research questions focus on the influences of virtuality and technologies on organizing and participation processes, and the resulting impacts on scientific outcomes. Investigating these questions required a research design for comparison across citizen science projects according to theoretical sampling criteria. A comparative case study design was chosen, with cases selected to reflect a combination of characteristics related to virtuality and technologies.

Case study strategies are common in applied fields such as information systems; as a research strategy, case studies focus on contextualized social activity and can involve multiple cases and levels of analysis (Yin, 1984). Case studies are generally considered most appropriate in the early stages of research on a topic, or to provide a novel perspective on an existing topic (Eisenhardt, 1989). The current study represents both, as research on citizen science as a phenomenon is currently at an early stage, and also takes a different perspective on the intersection of the existing topics of public participation in research, scientific collaboration, and online communities.

The initial theoretical framework for this comparative case study was developed in Chapter II. Concepts from inputs, states, processes, and outputs in the framework provided the primary concepts for data collection and analysis. Describing virtuality and technologies required collecting data on the current state of each project with respect to social and technological contexts. It was expected that these combined elements would influence organizing and participation processes. In turn, the processes' outputs and the processes themselves were expected to have an effect on the outputs of the project, particularly scientific knowledge and innovation.

Cases in comparative studies are selected to fill theoretical categories using replication logic, or to extend theory development with extreme cases. The case selection strategy for the study will be discussed in more detail later. Case studies collect multiple sources of data as evidence for triangulation of findings in a relatively flexible research process in which data collection and analysis occur concurrently (Perecman & Curran, 2006). These data are used to connect the research questions to the findings through a variety of analysis techniques, including within-case analysis and cross-case analysis. In the current study, several types of data to evaluate and further develop the initial conceptual framework were collected through field research methods, including interviews, participant observation, archival records, documentation, and artifacts. Data collection and analysis proceeded simultaneously, leading to written case descriptions representing within-case analysis and then to comparisons for cross-case analysis.

The knowledge gained from data collection and early stages of analysis for the first two cases shaped the theoretical sampling criteria for selecting the third case, and the iterative process served to further focus analysis as interesting themes emerged from the data. For example, both eBird and the Great Sunflower Project were reliant on information technologies from the start, making them fundamentally technology-enabled; one of the reasons to select

Mountain Watch as a complement was that it did not rely on information and communication technologies, which were later developed to extend the reach of the project. Emergent themes were integrated into the theoretical framework, and each case was re-examined after each such change. This iterative and concurrent process of data collection, analysis, and revision to the theoretical framework followed inductive analysis processes, combined with deductive analysis focused on refining the initial conceptual framework.

In summary, this research applied a comparative case study strategy to examine an emergent and non-deterministic phenomenon, using theoretical replication to compare the impacts of virtuality and technologies on citizen science project processes and outputs. Iterative data collection, analysis, and theory development stages of research were conducted simultaneously. Case selection criteria and brief descriptions of case study sites are the subject of the next section.

3.2 Case Selection

Careful selection according to theoretical sampling criteria is critical to addressing the research questions in a comparative case study. Typically each case will represent a different variation on the sampling criteria. The diversity of citizen science projects across each of the potential dimensions for sampling suggests that commonalities identified in such diverse settings are likely to be particularly interesting and useful for understanding the shared challenges and emergent solutions in the domain of citizen science organizing practices.

The case selection for the study focused on projects that are organized around primarily scientific goals and involve participants in collecting observational data about the natural world. The majority of citizen science projects follow this general model. Choosing projects in this (broad) category means that findings have greater potential to be applicable to the widest possible range of citizen science projects. The sampling strategy attempted to be

as representative as possible while making no claims as to the degree of applicability to the larger population of citizen science projects, as theoretical sampling is not equivalent to statistical sampling. Because these projects focus on data collection through observation the fundamental participation tasks are similar, which supports meaningful comparisons. Theoretical sampling criteria, discussed next, were based on the conceptual framework, which provided additional guidance for case selection.

3.2.1 Theoretical Sampling

Theoretical sampling with replication logic means strategically choosing cases that provide theoretically-based contrasts. The project-level inputs from the initial conceptual framework (purpose, community, environment, technologies) provided useful and straightforward criteria for case selection. Importantly, project-level inputs are among the most readily identifiable and distinguishable characteristics of citizen science projects, and are sufficiently transparent and accessible enough to enable an informed selection without extensive prior engagement with the project. For example, the community that forms the target audience for a citizen science project is often readily identified by the focus of the participation activities and research domain, e.g., a project focused on collecting data about birds would have natural appeal to birders and birdwatchers. The goal for case selection therefore was to identify several cases with suitable variation in their purpose, community, environment, and technologies that could adequately span the breadth of the larger domain of citizen science practice in each of these categories.

Purpose Two primary dimensions of purpose were considered for sampling: these were 1) the scientific interests and goals that comprise the “science” focus of citizen science, and 2) the mission of the project with respect to broader goals. In organizational and institutional contexts, mission is congruent with the guiding principles that both help describe the purpose

of the organization, institution, or project. Mission also provides a tool for decision-making with respect to organizing. The missions of most citizen science projects and the larger institutions with which they are often associated are frequently within the same subset of broader goals, particularly for observation-based citizen science, so broad variability on this dimension of the concept of purpose was less important for sampling. If anything, a common focus on research, education, and conservation would make the cases more representative of the majority of the citizen science project population.

Instead, variation along the dimension of scientific interests is more important: the projects selected for comparison should be in different scientific domains, with different research goals. Therefore, not all of the projects could be in ornithology, because birders are different from other citizen science participants. Likewise, not all of the projects should focus on questions of species abundance and distribution, for example, because their activities might be too similar for useful comparison; rather, they should represent different types of research questions.

Community As interpreted in the theoretical framework and applied to sampling, community means that the cases need to draw on different communities of contributors. In citizen science, community is typically congruent with scientific domain variability. The intuitive expectation is that people who self-select for participation in citizen science projects focusing on birds come from a community of birdwatchers and birders, and those who help with trailside invasive species monitoring are typically members of a hiking community. Although there are frequently members of other communities and sub-communities that citizen science projects may engage as participants, community as used in this sense refers to a community of practice—experienced hikers, for example, have common practices such as following the Leave No Trace principles of outdoor ethics to minimize environmental impact, and share

numerous strategies to reduce pack weight and eat well on the trail.

Environment The concept of environment was broadly defined in the theoretical framework. For the purposes of sampling, environment refers to organizational contexts and/or the broader organizational field in which a project is situated. Therefore, ideal projects for a cross-case analysis would have varied forms of organizational and institutional support and constraints, because these factors affect project resources and purpose. Such variability may appear in the nature of the institutional arrangements. For example, a project may be the product of a single nonprofit organization or a network of organizations.

Another form of variation for the organizational context of sociocultural environment has to do with the types of institutions involved. Citizen science projects are most frequently operated by academic researchers and public-sector groups, namely nonprofits and governmental agencies such as the U.S. Department of Agriculture, National Park Service, or individual states' departments of environmental conservation or protection. In terms of optimizing data collection for breadth, a combination of these forms of environmental variability is preferable to single-dimension variability. Examining projects operated only by governmental agencies with different institutional arrangements, or projects organized by an individual organization across nonprofit, academic, and governmental contexts, would be valuable but less representative of the broader population of projects.

An additional point of comparison related to the project's operating environment is the resources that can be brought to the project. These include staffing and fiscal resources; larger staff and budgets are related to several other project characteristics, although causality is not entirely clear. However, findings from Wiggins and Crowston (2012) suggested that more staffing may mean more participants can be supported, higher annual budgets may yield more sophisticated technologies and more extensive outreach, and in combination,

greater resources are likely to lead to higher volumes of data outputs and scientific outcomes. The degree to which these generalizations may be true is expected to be dependent on many factors, such as the project goals, complexity of participation processes, and the way contributions are measured, which are often incommensurate across projects.

Technologies The technologies in use for each project are a point of contrast that is directly related to the research questions. There are numerous facets of technologies that could be employed as a basis for theoretical sampling. One point of contrast relates to the technologies used for making, managing, and reporting field observations, as these are the common core tasks across the target population of citizen science projects. For example, paper-based record making in the field is very common in observational citizen science. The nature of these uses of paper, however, can differ substantially, from a protocol-based data sheet to multiple types of species lists which may follow established community conventions (see Figure 3.1a) or may be generated according to individual field observation habits (see Figure 3.1b.) The ways that these material technologies are implemented to support project participation were an important dimension for sampling.

The nature of information and technologies that support participation is another aspect of the sampling criterion. The specific qualities of ICT can be difficult to meaningfully evaluate at an adequately general level as to be of use in analysis. The heuristics used to streamline this distinction were the overall degree of information technology sophistication, which was evaluated according to the fundamental type of web-based technologies (e.g., devoted purpose-built platforms, content management systems alone, or standard websites with data submission forms) and the availability and breadth of means for data access by participants. While it is a substantial simplification of the complex variability of information technologies that can be used to support public participation, these interrelated characteristics have been

observed in prior interactions with citizen science projects as a key feature that seems to make a substantial impact on participation and organizing processes, which are the focus of the research questions. The cases selected according to these sampling criteria are described next.

3.2.2 Selected Cases

The cases selected by the theoretical sampling strategy discussed in the prior section are briefly introduced here: Mountain Watch, the Great Sunflower Project, and eBird. These case studies maximize the depth of the research based on intensive data collection for one project and complementary but less intensive data collection for two others. For a number of reasons, there was simply more data available for eBird, for which a substantial volume and variety of data were collected. Therefore, this case formed the initial basis for schema development during later analysis. The Great Sunflower Project and Mountain Watch complemented eBird (and one another) by offering theoretical and experiential counterpoints.

Criterion		Mountain Watch	Great Sunflower	eBird
Purpose	Mission	Conservation, education, recreation	Research, education	Research, education, conservation
	Scientific Interests	Climate change effects on alpine habitats	Plant-bee relationships	Bird abundance & distribution
Intended Community		Hikers	Gardeners	Birders
Environment	Institutions	Single nonprofit	Academic	Nonprofit partnership
	Resources	1.5 FTEs, \$15K	0.5 FTE, \$13K	4.5 FTEs, \$300K
Technologies	Paper	Structured data sheet	Structured data sheet	Variable, optional
	Digital	Organization web-site section	Open source CMS	Purpose-built software system
	Data access	Limited	Very limited	Extensive

Table 3.1: Application of theoretical sampling criteria to selected cases.

As the conceptually-focused comparisons between the cases are the substance of the following chapters, the contrasts used to guide sampling are summarized in Table 3.1. In the

table, mission represents explicit organizational mission or goals expressed as project mission. Resources are summarized in full-time employee equivalents (FTEs) and approximate annual operating budget. The remainder of this section provides a brief overview of the cases.

Mountain Watch

Mountain Watch is a two-part citizen science project designed and operated by the Research department of the Appalachian Mountain Club (AMC), a membership-based trail club whose mission is to support conservation, education, and recreation in the northeastern mountain ranges of the Appalachian ridge. Mountain Watch enlists hikers in evaluating air quality through visibility measurement and in collecting observations of flowering plants for climate change research. The project is geospatially constrained, collecting data primarily in the White Mountains of New Hampshire, the largest alpine region in the Northeast U.S and home of the AMC's main visitor centers, administrative offices, and backcountry facilities (the "High Huts"), which are operated under a special use permit from the U.S. Forest Service. The flowering plant project, now the primary focus for participation, gathers long-term data to monitor the effects of climate change on fragile alpine ecosystems by examining the timing of plant life cycle stages (phenology), such as flowering and fruiting (phenophases). Although hikers can report data online for any location in the broader northeastern U.S. region where the target plants are found, the project's primary participation comes from hut guests.

Starting at an AMC facility, hikers pick up a packet with an instruction and identification guide, data sheet customized to location, and pencil in a plastic zip bag. Hikers locate monitoring plots using provided maps and text descriptions, and indicate which species are present and whether they are in any of the indicated phenophases (e.g., "before flowering," "flowering," "after flowering"). The completed data sheets are then dropped off in collec-

tion boxes at any of the eight huts or the visitor centers at Pinkham Notch and Crawford Notch. Mountain Watch is a mature project that has methodically fine-tuned its participation protocol over a period of several years to produce increasingly scientifically useful data. Mountain Watch slowly transitioned from analog to digital support for data collection, and has fully incorporated the project into the daily operations of its backcountry facilities.

The Great Sunflower Project

As previously described in Section 1.1.2, the Great Sunflower Project (GSP) focuses on pollinator service, that is, bee pollination activity. Participation in the GSP is mostly performed by independent volunteers, although it can be done in small groups. Volunteers plant Lemon Queen sunflowers (other flowering plants have been added through program expansion) and can optionally track the plant's growth progress by reporting phenology observations while they wait for their sunflowers to grow.

Once the sunflowers bloom, participants choose a flower that is in the appropriate stage of development to attract bees, and describe the observation conditions. Next, they observe the selected bloom for fifteen minutes, recording the times at which bees visit, and attempt to identify (and optionally photograph) the visiting bees. The majority of data are then entered by the volunteers into an online database, although some participants submit paper observation forms by postal mail. The primary focus of the project is collecting data for scientific research on pollinator service at a national scale, as it is an important indicator of local ecological health. To date the project has been very successful in attracting volunteers, although the scientific outcomes are not as yet evident due to revisions to data collection procedures. The Great Sunflower Project represents a young, underfunded, and technologically disadvantaged citizen science project that has shown remarkable potential and resilience despite substantial challenges.

eBird

eBird is a popular citizen science project developed by the Cornell Lab of Ornithology (Sullivan et al., 2009), a leading organization in the development of citizen science practice.

eBird allows birders to keep birding observation records online:

“A real-time, online checklist program, eBird has revolutionized the way that the birding community reports and accesses information about birds. Launched in 2002 by the Cornell Lab of Ornithology and National Audubon Society, eBird provides rich data sources for basic information on bird abundance and distribution at a variety of spatial and temporal scales.

eBird’s goal is to maximize the utility and accessibility of the vast numbers of bird observations made each year by recreational and professional bird watchers. It is amassing one of the largest and fastest growing biodiversity data resources in existence.”¹

eBird contributors can submit basic data by completing online checklists of birds seen and heard while birding. The system also provides tools suited to supporting independent inquiry. Users can query and visualize their own data and that of others, exploring interactive maps, graphs and charts. Contributed data are aggregated, reviewed by local experts for quality when flagged by automated data filters, and then integrated into the Avian Knowledge Network, a public archive of observational data on bird populations across the Western hemisphere. eBird represents a mature, well supported, and technologically sophisticated project that is engaging volunteers internationally on a massive scale, receiving up to three million observations monthly by 2011.

3.2.3 Comparison of Selected Cases to Broader Population

A survey of citizen science project organizers conducted in 2011 focused on project characteristics, and helped establish the representativeness of the cases. A description of the survey methods and instruments are included in Appendices B, C, and D. Several details

¹<http://ebird.org/content/ebird/about>

Feature	Survey	Mountain Watch	Great Sunflower	eBird
Annual budget	\$100K (average), \$35K (median), \$20K (mode)	\$15K	\$13K	\$300K
Staffing	1–1.5 FTE	1.5 FTE	0.5 FTE	4.5 FTE
Age in years, 2012	14 (average), 10 (median), 3 (mode)	8	4	10
Funding sources	Grants & donations (monetary, in-kind); average 5 sources	Multiple grants, donations	Multiple grants, donations, merchandise, referral sales	Multiple grants, donations, sponsorships, service fees
ICT	Website, email, social media, publications, articles, graphs & charts, data summaries	Website, email, interactive map	Website, email, map, summary data, social media	Website, email, multiple types of maps, on-demand reports, charts & graphs, social media, articles, blog, data tools, publications
Data quality	Expert review, photos, data sheets, replication, training programs	Expert review, data sheets, replications	Expert review, data sheets	Automatic filtering, expert review, photos & replication

Table 3.2: Summary of average survey respondent project characteristics and cases.

from the survey are presented here to contextualize the case selection and later discussion, summarized in Table 3.2.

To better understand the resources that projects are able to devote to various aspects of development, implementation, and improvement, the survey asked about levels of staffing, annual budgets, and sources of funding. Responding projects had between zero and over 50 paid full-time equivalent employees (FTEs), with the majority of projects employing 1–1.5 FTEs. All of the cases in the study were fairly representative of the usual staffing arrangements. Annual budgets ranged from \$125 to \$1,000,000 (USD or equivalent), with an average of \$104,882 but with a median of \$35,000 and a mode of \$20,000. The cases in the study had budgets that ranged from approximately \$13,000 to \$300,000 per year, representing both the high and low ends of the distribution for fiscal resources. Responding projects were widely variable with respect to the age or duration of the project. A few were not yet operational, and one was over 100 years old, but most were started in the last 10

years. The cases for this study included projects that were 10, 8, and 4 years old at the time of data collection, making them representative of the larger population.

The survey asked organizers about the types of funding sources that they use to support their projects. Most projects relied primarily on federal or other grants, followed by in-kind contributions and private donations. Organizers leveraged up to five different funding sources to meet their expenses. The cases chosen for this study had a range of funding sources, primarily grants, both monetary and in-kind donations, sponsorships, service fees, merchandise sales, and in-kind contributions; each project used a range of resources that are best suited and available to the organizers.

To learn about the range of information and communication technologies supporting citizen science, the survey asked about tools in current use for communication with project participants. Several technologies were used for communication among project organizers and between organizers and participants, with websites and email being the most common by a large margin. Other common communication tools were print publications, research articles, and several types of data representations, including maps, graphs, charts, and data querying and summary tools. As later discussion will demonstrate, the differences between the websites for eBird, the Great Sunflower Project, and Mountain Watch are difficult to adequately summarize and compare to the details reported on the survey. Nonetheless, the cases selected for comparison have implemented many of the same types of technologies adopted by the broader population of citizen science projects.

The majority of responding project organizers employed multiple mechanisms to ensure data quality, for which specific requirements were dependent on the goals of the project. Data quality typically refers to the precision, accuracy, and reliability of data contributed by participants, which affect the utility of the data for scientific research differentially based on research goals, analysis methods, and methodological standards. The most common combi-

nations reported included expert review along with additional documentation of observations. The frequency with which expert review and additional documentation is employed in citizen science reflects in part the dominance of data collection as the primary task for contributors, but also concerns over accurate identification, for example, of species or phenophases. For the cases in this study, eBird employed automatic filtering and expert review; Mountain Watch used paper data sheets, expert review, and multiple replications of observations; and the Great Sunflower Project relied primarily on expert review and data sheets. The cases are a reasonable representation of the larger trends in data validation practices.

For these descriptive characteristics of citizen science projects, the cases selected for the study were typical with respect to staffing, age, sources of funding, types of technologies used, and data validation mechanisms. They also included two projects with budgets close to the mode and one with above-average fiscal and staffing resources, providing opportunity to learn how variations in resources can influence project development.

3.3 Case Study Data Collection and Analysis

Contextually-focused data collection requires attention to the people and environment of a social phenomenon. Multiple sources of data were collected to produce a record that provided a broad view of the phenomenon of interest and allowed for triangulation of findings. The primary data sources were interviews with project organizers in a variety of roles, participant observation, and documents.

Data collection and analysis were guided by the research questions and the conceptual framework discussed in Chapter II, which acted as a sensitizing device. Data analysis was conducted concurrently with data collection, which guided further data collection with the developing understanding of the phenomenon. This section focuses on the research process, types of data collected, and analysis procedures.

3.3.1 Research Process

Case selection, negotiating access, and seasonality for participation and observation dictated the timelines for completing the research. Figure 3.2 depicts the research process as it unfolded, with overlapping data collection and analysis processes.

Data collection relevant to these cases began in 2009 with initial interviews for a pilot study that included organizers of the Great Sunflower Project and Mountain Watch, with additional subsequent interviews in fall 2010 and summer 2011 for these projects. For eBird, data collection began with a large number of internal documents and interviews with project organizers over a period of four months in the summer and fall of 2010. Participation in both eBird and the Great Sunflower Project lasted more than a year.

Initial analysis of interviews from eBird started in December of 2010, and led to the concretization of the case selection criteria. Mountain Watch was formally recruited as a case in January of 2011, although participant observation data collection was delayed until June due to the seasonal aspects of participation.

Coding of interview transcripts resumed in July and August of 2011, using a revision of the conceptual framework as a focus for analysis. Coding started with the eBird case, and then continued to the Great Sunflower Project and Mountain Watch; the coding process was repeated in multiple analysis cycles. At the same time, initial case descriptions were written for each project, and interview transcripts were sent to each interviewee for verification. Again, additional revisions were made to the theoretical framework as the writing process provided further focus to the analysis. This iterative process highlighted differences between each case; the findings from the eBird case were moderated by the other two cases, leading to additional revisions to the theoretical framework.

In the fall and winter of 2011, continuing analysis overlapped with writing, leading to completed case descriptions, structured around the version of the theoretical framework

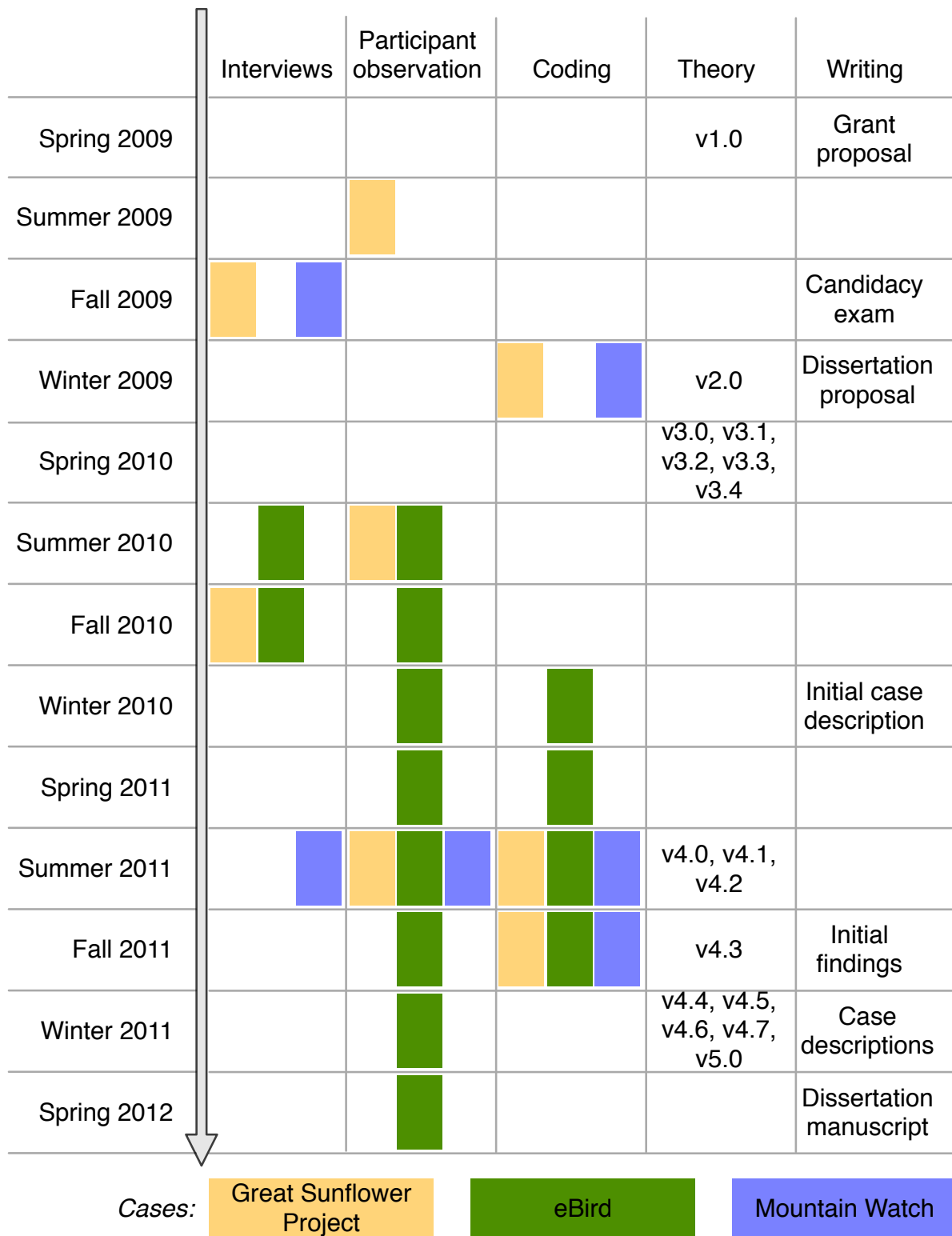


Figure 3.2: Research timeline for data collection, analysis, theory revisions, and writing.

presented here. Throughout the process, changes to the theoretical framework were reviewed and discussed with peers. The remainder of this section discusses the various types of data collected for the study and the analysis processes in further detail.

As Eisenhardt (1989) discusses, the issue of closure in case study research is dual; researchers must determine when to stop adding cases and also when to stop iterating between data and theory. The concept of theoretical saturation advocates continued sampling of cases until there are no new insights. Three cases were selected to address the research questions through theoretical replication, with each case representing a different variation on the sampling criteria. These choices were based on theoretical criteria but also reflect choices based on deep familiarity with the goals of the study, the context of the research, and the types of data that were considered most useful for investigating the research questions.

The second form of closure mentioned by Eisenhardt (1989) refers to iteration between data and theory, which should stop when there is minimal incremental improvement to theory. This goal was moderated with consideration for resources and the planned timeline for completion. Changes to the theoretical framework decreased as analysis continued until there were no additional modifications that were warranted by the data from these cases.

3.3.2 Data

Data collection proceeded on a case-by-case basis, as each case was subject to different constraints and provided different opportunities for data collection. Similar types of data were collected for each case, subject to availability, and are more thoroughly described in the case study descriptions. The case study methodology took advantage of the flexibility of qualitative field research methods, combining a variety of elicitation methods to build a deep understanding of the cases (Perecman & Curran, 2006; Bailey, 2007). The overall case study data collection approach shared many of the characteristics of traditional ethnography, including negotiation of access, long-term participation and observation, longitudinal

interviews, and writing of field notes (Spradley, 1979; Fetterman, 1998). Similarly, analytic practices included ongoing memoing, coding, and description.

Throughout the research process, additional involvement with the citizen science organizer community and thematically related research projects provided additional insight into the comparability of these cases to the larger population of citizen science projects. Engagement with the practitioner community was nontrivial, and is described further later in this section.

This involvement provided substantial contextual background extending beyond the case study sites themselves and into the larger organizational field. These experiences, while not quantified or quoted in the analysis, were an important part of developing a theoretical framework that is expected to apply to a much broader range of citizen science projects than only those studied intensively in the current research. When broad statements about citizen science projects are made in the case descriptions and analysis, these assertions are made based on the wider range of experiences with the citizen science practitioner community and prior empirical research on citizen science.

Interviews

Interviews elicit the firsthand accounts of the people involved in the phenomenon; they are employed in various forms of qualitative and quantitative study. Multiple interview types are typically employed with field research approaches, including informal, ethnographic, semi-structured, and formal interviews (Kvale & Brinkmann, 2008). This research included primarily semi-structured and informal interviews, with some longitudinal interviews that provided deeper insight into project development. Due to the wide variability of interviewee roles with respect to each of these cases, several variations on the initial semi-structured interview protocol were used as the research evolved and for different interviewees based on their relationship to the project (see Appendix A for a representative interview protocol).

The interviews elicited narratives of project development that were similar in some re-

spects to an oral history, with each account reflecting the experiences and perspectives of the participants. Semi-structured interviews with project leaders and other organizers provided complementary perspectives for a more holistic understanding of the cases. Whenever possible, interviews were recorded and transcribed, and all interviews were augmented with field notes.

Participant Observation

Participant observation is a field research data collection method for developing the deep understanding the context and practices of a social phenomenon that generally requires “being there” (Spradley, 1980). In this discussion of research methods, the use of the terms *participants*, *participation*, and *observation* may become unclear. Participants are the people contributing to the project through participation processes, and these processes are a concept in the research questions and conceptual framework. In addition, contribution to these projects involves collecting observations of natural phenomena. The researcher, as a participant observer, both participates and observes participation of others. For this study, I was a participant in the larger organizer community and a participant the case study projects (in the same sense as other contributors). As contributing participant, I contributed observation data to the projects, as do other participants. In the process, I also observed other participants who made observations of birds, bees, and flowers. While potentially confusing in reference to participant observation methods, the terminology is consistently applied in subsequent chapters: *participants* are contributors, *participation* is the process of contributing, and *observations* are the data that participants contribute.

Field research provided an opportunity to develop a deep, experiential understanding of the participation processes in context. During participant observation, field notes provided data documenting the experiences and developing understanding of the researcher (Emerson, Fretz, & Shaw, 1995). Participant observation took two forms, both as a contributor in each

project, and in citizen science organizer meetings and community events. With respect to participation with organizers, extensive experience with the broader context of citizen science was gained through attendance at ecological conferences, stakeholder meetings, invitational workshops, NSF grant advisory committees, DataONE working group meetings, project planning committee meetings and teleconferences, coauthorship of ecology journal articles with project organizers, and a pilot study focused on the development of a regional network of citizen science projects. Therefore, comparisons drawn to other citizen science projects are based on four years of immersive experiences working with these groups.

Participant observation also included participation as a contributor in each project, although at varying levels of intensity congruent with the nature of each case and my inherent interest in the activities. Participation also included reading and posting to email listservs and online forums from the standpoint of a non-researcher (Best & Krueger, 2004; Hine, 2000; Ruhleder, 2000). Interactions recorded via electronic means are a form of secondary data, but as my participation was concurrent with the generation of most of the records consulted during analysis, they are not separately categorized as such.

The duration of participation varied by the project structure: eBird participation lasted well over a year (continued in large part due to the enjoyable nature of the activity), while Great Sunflower Project participation was limited to summers over three years and Mountain Watch participation was constrained to a single very intensive week. The locations of participation were also disparate. While the majority of eBird participation took place in my back yard, it also occurred in 140 other distinct locations in 19 states and provinces across four countries. The Great Sunflower Project participation was limited to my front yard, where I grew sunflowers for observing bees. Mountain Watch participation occurred along hiking trails in a relatively small area of the White Mountains of New Hampshire surrounding the Appalachian Trail. Each form of participation required different skills and

knowledge, and I learned a great deal from the domain-specific content of each project in addition to learning about the projects themselves.

Observation of other participants' experiences and behavior was primarily conducted via email lists and forums for two cases (eBird and the Great Sunflower Project.) Indirect and virtual observation of contributors' practices was aligned with the experiences of other participants, who also participate remotely, and provided substantial insight into the norms, interests, skills, and values of participants, in addition to substantiating many of the claims that organizers made about project participants. The nature of participation in Mountain Watch afforded more direct opportunity for observation, which enhanced my understanding of participation more generally. Notably, however, the nature of the Mountain Watch project also made direct observation of other participants much more relevant than it would have been in the other cases. In eBird and Great Sunflower Project participation is primarily undertaken by individuals (or in small groups), whereas Mountain Watch participation often includes substantially more direct in-person interaction with organizers. Two different modes of participation in Mountain Watch were undertaken to provide additional context for the variations on the experience as relates to independent versus guided group participation.

Documents and Artifacts

Both documents and artifacts can provide rich sources of data. Documents provide a view of interaction within a social group, and often play an important role in case study research (Yin, 1984). The specific types of documents collected for each case are detailed further in the case descriptions. They included several hundred individual documents, such as grant applications, protocols, data sheets, promotional materials, newsletters, scholarly articles, and numerous photographs made during participation. These documents further clarified project history, processes, and outputs.

These data sources were used in two ways: to provide background and context for the case

studies, and as sources of data for triangulation of claims made in interviews and for assessing the commonality of participation experiences from participant observation. As a form of background to the case studies, documents such as webpages and grant applications provided additional information on project histories and goals, organizational context, and insight into how organizers communicate with participants. These data sources also supplied additional detail on participation procedures, how to use supporting technologies, documentation of the places in which participation occurs, and evidence of the scientific outcomes of the projects.

As mentioned above, the posts on listservs and forums were used to verify organizer claims, e.g., that participants encountered problems using data entry forms, or found website features motivating and exciting. The opinions and positions expressed by project participants were also compared to those of the researcher to better understand the likelihood that these experiences were shared by others, and therefore potentially relevant to the broader population of participants.

3.3.3 Analysis

As previously mentioned, data analysis was an ongoing, iterative process that began in late 2010 and continued through the fall of 2011. The analytic process adopted for this study was a combination of deductive and inductive coding of interview transcripts. Coding was conducted using the TAMS Analyzer open source software for Mac OS X (Weinstein, 2012).

Analysis Processes

The analysis was initially guided by an early version of the theoretical framework (Wiggins & Crowston, 2010), employed in a deductive fashion, followed by iterative inductive analysis. The initial coding focused on only one case (eBird) during the development of a schema that highlighted themes related to the initial conceptual framework, as well as new themes that emerged through analysis. This initial analytic phase also helped to identify pertinent and

rich passages of data to use for further study.

Next, inductive open coding within the text identified as relevant to the themes from the framework yielded a much broader set of concepts more directly related to the research questions than the initial framework. The process highlighted several recurring and important concepts that were not included in the earlier theoretical framework. These were subsequently incorporated into the framework and used as a coding schema for the full set of interview transcripts. Many elements of the initial framework were revised to better reflect concepts derived directly from the empirical data.

After revision of the theoretical framework, the concepts derived from the first case were used to deductively code the entire set of transcripts for all three cases. As the second and third cases were coded, additional themes emerged, and were again incorporated into the coding schema. The previously coded transcripts were reviewed for evidence of these themes, and analytical memos were produced to track the evolution of the theoretical framework as well as insights produced in the process of coding transcripts. Once all transcripts had been coded with the concepts from this version of the theoretical framework, the coded text was systematically retrieved on a case-by-case basis and summarized in memos focused on each theme.

Because the coding was conducted at the level of the thematic unit (variable chunks of relevant text), there were numerous overlapping codes that indicated conceptual relationships. These co-coding instances were automatically identified with software reports, allowing easy examination of some of the relationships between concepts for both the individual cases and combined cases. After examining these instances, those for which two codes coincided frequently were noted, although there was no assumption of causality in these relationships. Diagrams for each individual case and the three cases together were then generated to visualize the relationships between each concept, with connections weighted by relative frequency

of the code overlaps, and concepts weighted by their overall frequency of occurrence.

These visualizations helped guide the descriptive analysis of each case by focusing attention on the prevalent conceptual relationships. For example, while the high frequency of technologies as a theme in the eBird case was due in part to the interview sampling, it was also reflective of the central role that the eBird technical systems play in the participation processes and scientific outcomes; these links were therefore discussed in some depth in the case description. Rich process models of participation processes were also produced for each case (see example in Appendix E), which helped direct exploration of variability in the skills and resources needed to support participation and assisted in identification of critical differences in these processes for each case (Jensen & Scacchi, 2005).

The case study analyses were developed from both descriptive data and the coded interview transcripts. Each case was written to provide background information about the project history, organizational context, participation processes, and supporting technologies. The coded interview transcripts were then mined for quotes and examples that demonstrated specific theoretical concepts or relationships between concepts. These quotes were then organized to structure the analysis of each case around the concepts from the theoretical framework and the research questions. The relevant theoretical concepts are specified in the titles of many subsections in the case study chapters.

The emergent findings reported in Chapter VIII were generated through an iterative inductive process. A series of themes connecting the concepts in the theoretical framework and highlighting commonalities between the cases was developed through review of the interview transcripts, visualizations of coding, and comparison of the evidence for the theoretical concepts in each case. Memos made during review of each individual theoretical concept tracked the relationships between the concepts, noted details of the projects that were taken for granted by interviewees, and identified repeated instances of these relationships across

the cases.

These processes yielded an initial set of seventeen topics. The list of topics was then condensed into the resulting five themes discussed here by consolidating related topics into coherent themes. With supporting evidence from the interviews, which made both explicit and implicit connections between the theoretical concepts, the themes were further developed to demonstrate the interactions of these concepts with the context of practice. In addition, these emergent themes were used to address the answers to the research questions in Chapter IX.

Finally, the writing process stimulated further analysis by narrowing the review of concepts and their relationships, which helped highlight additional points of comparison between the cases. The process of writing the case descriptions generated new insights into the distinctive features of each case as well as the reasons for the differences between cases, which were captured in both the writing of the case descriptions and the cross-case analysis.

Writing Conventions

Several writing conventions are used in this manuscript; they are described here for clarity. All identities are anonymized with portions of the Latin names of organisms related to the scientific domain of study. For example, the pseudonym Clintonia is derived from *Clintonia borealis*, Bluebead lily, which is one of the forest flower monitoring species in Mountain Watch. The exception to the rule is in the case of project founders whose full names are used because the public nature of information about their roles in the respective projects means that anonymization provides no identity protection.

There are a variety of roles that individuals take in citizen science projects; they are reflective of the structure of the projects and essentially self-assigned. These roles described according to the terminology used by interviewees, with additional specification to clarify distinctions based on organizational membership. Several terms are used consistently

throughout to describe these roles. *Project leaders* are in charge of managing the project. *Project organizer, team member, or project staff* are terms referring to other staff who contribute to the project. The distinction between project leaders and project organizers is most relevant for the eBird case study, as there were a wider variety of roles in that project than the others; these terms should be regarded as synonymous for the other cases. *Partner project organizers* take a leadership role in external organizations that explicitly support the project, but are not officially part of the project leadership team. *Participant* and *contributor* refer to the citizen scientists who contribute to and participate in the project as volunteers. Occasionally the term *project member* is used to refer to contributors as well, but is always distinct from team member. *Registrant* refers to people who have indicated interest or willingness to participate in the project, typically by creating an online account; these individuals may or may not be contributors, which is relevant primarily in the case of the Great Sunflower Project. *Interviewee* refers to those individuals who provided interviews for the study; for these cases, the term participants never refers to interviewees.

Quotations from interviews are formatted in APA style. The numeric references following citations refer to the location of the character range within the interview transcript wherein the cited text can be found, e.g., (Pinicola, 1234–5678). For practical reasons, these text locations are not precise to the exact wording, but always include the full text cited. The text locations refer to the raw transcripts, i.e., additional characters added by coding are not included in these character counts. This ensures that if the raw transcripts were examined, the precise location of quotations could be found.

All quoted text is verbatim, subject to minor omissions in transcription that do not affect meaning in the analysis. Non-word utterances were not transcribed, nor were sequentially repeated words, e.g., “I found that, that it, that it just confused things” would be transcribed as “I found that it just confused things.” The phrase “you know” was omitted when used

habitually in the same fashion as a non-word utterance *and* clearly not related to the content of the statement; it was included when the relationship of the phrase to content was either contextually relevant or unclear. Although this required a judgment call, reviewing the audio recordings of each interview made these choices straightforward; the interpretation of the use of this phrase was unchanged from the original context of spoken conversation. These omissions were made for the sake of clarity and brevity, as the vernacular speech of some individuals would otherwise require ellipses so frequently that quotations would become difficult to read.

3.4 Research Quality

As with any research design, the case study methodology has its limitations. Assessing the quality of qualitative, interpretive, contextualized research poses different challenges than quantitative research, for which statistical tests provide established means to evaluate validity. Contextually-grounded research aspires to analytical or theoretical generalizability, the ability to apply the theoretical insights derived from one context to others, which is particularly relevant for contributing to practice as well as theory. The limitations of any given methodological approach are most often criticized with respect to validation of the research, which will be discussed next, along with the validation strategies used in this study.

Modes of validation depend on the purpose of the study and the researcher's philosophical position on objectivity and ontology. Intersubjective evaluation of freedom from bias in qualitative research involves examining the findings by degree of validation, theoretical perspective, reflexivity, articulation of bias, and the case for generalizability. In addition, explicit discussion of the researcher's positionality, or personal qualities and values with respect to the research, allows others to understand the ways in which the researcher's individual

personal characteristics, context, and knowledge affect her experiences and findings; these details are discussed below. Positionality is just one aspect of the reflexivity documented in the written report, and awareness of the impact of the researcher's perspective on the research must be maintained throughout the research process (Davies, 2008). The remainder of this section further discusses validation strategies and positionality.

3.4.1 Validation Strategies

Case study research employs multiple tactics for addressing research quality at each stage of research. According to Yin (1984), four aspects of research quality that case study research designs must address are construct validity, internal validity, external validity (generalizability), and reliability. Lincoln and Guba (1985) argue, however, that these criteria drawn from quantitative research frameworks do not translate well to qualitative research. Instead, they suggest the criteria of credibility, transferability, dependability, and confirmability. These four criteria were used as the guiding principles for supporting research quality.

Credibility equates to internal validity, establishing that the research results are believable from the perspective of participants in the research. Transferability takes the place of external validity, and is established through description of research contexts and assumptions so that readers can judge whether results are transferable to other contexts. In addition, the comparative case study design with theoretical replication logic improves transferability (Yin, 1984), and supports theoretical generalizability. Dependability departs from reliability in acknowledging that contextualized social phenomena cannot be measured twice. Instead, the researcher must describe changes to the research context and how the research approach was affected by them. For example, this study incorporated dependability checks through creation and peer review of an audit trail, two stages of validity verification through participant review, and periodic peer review with outside experts. Since the criterion of objectivity is contrary to the subjectivity of qualitative research, confirmability focuses instead

Criterion	Strategy
Credibility	Multiple peer and external reviews, participant review of transcripts and case study descriptions
Transferability	Comparative design with replication logic
Dependability	Audit trail with peer review, memos throughout data collection and analysis
Confirmability	Triangulation, audit trail, documentation of analysis and theory development

Table 3.3: Validation criteria and strategies.

on whether the results can be confirmed by others. Confirmability is a point of evaluation achieved through documentation of the research process, peer review, and triangulation of data sources. The strategies employed to address these four criteria for evaluation in the current study are shown in Table 3.3 and discussed in the remainder of this section.

As is often recommended for field research, this research study involved creating an audit trail documenting research decisions to help maintain research reliability and internal validity. The audit trail documented the connections from the evidence in multiple sources of data to the analysis process, and the chain of reasoning leading to the interpretation. Experts in the domain of citizen science practices at the Cornell Lab of Ornithology were consulted to verify that the emergent findings were not an artifact of sampling, but could reasonably apply to other citizen science projects.

To strengthen credibility and dependability, the findings and interpretation were also subject to review by the participants at two stages, in addition to ongoing verification from contact with key informants. Initial participant review came from sending a copy of interview transcripts to each participant for examination and modification as desired; only one interviewee opted to make any changes to an interview transcript (removal of a few lines of speculative commentary.) As case descriptions were completed, key informants from each case site were asked to review the completed chapter for their project and provide corrections. These reviews provided opportunity to verify factual accuracy as well as interpretation on a per-case basis. Each complimented the depth of the description and suggested minor

Principle	Definition	Strategy
Hermeneutic circle	Iteration between meanings of interdependent parts and whole	Multiple iterations through analysis and theory development; evolving data collection
Contextualization	Critical reflection of research setting demonstrating emergence of current situation	Organizational context and project history background documented for each case
Interaction between researchers and subjects	Critical reflection on social construction of data through interaction	Analysis acknowledges potential influence of researcher focus on interview data
Abstraction and generalization	Relating findings to other theories and concepts	Synthesis of other conceptual models into theoretical framework
Dialogical reasoning	Sensitivity to contradictions between theoretical preconceptions and actual findings	Iterative theory development process; discussion of additional emergent themes
Multiple interpretations	Sensitivity to possible difference in interpretations among participants	Development of project description through triangulation
Suspicion	Sensitivity to possible biases and systematic distortions in participant narratives	Elicitation of accounts from both within and outside of project leadership

Table 3.4: Evidence for conformity to evaluation principles for interpretive field research.

corrections. The case descriptions were also peer-reviewed to verify theoretical coherence and interpretations of the evidence.

In addition, Klein and Myers (1999) summarized seven principles for conducting and evaluating interpretive field research. This research was checked and adjusted according to these principles to support the quality of the research process and product. The seven principles are summarized in Table 3.4, along with strategies for and evidence of their use. Notably, even the finest researchers do not typically meet every principle in a single study, as discussed by Klein and Myers (1999). These principles are upheld to varying degrees, which is true of the current study as well.

In summary, the strategies employed to address research quality included the use of replication logic in the sampling design, creation of an audit trail for review by colleagues, and multiple reviews of both data and findings by peers, participants, and expert practitioners. The combination of techniques drew upon the strengths of different audiences to respond to different aspects of research quality throughout the research process. In addition, the

following discussion of researcher positionality reveals the personal characteristics and values of the researcher which may influence the quality and interpretation of the research.

3.4.2 Researcher Positionality

The role of the researcher is critical in qualitative research, and becomes all the more evident in field research involving participant observation. The researcher has substantial influence on the way the research is conducted and the findings that are generated. Acknowledging that neither research nor researchers are value-free, this section briefly discusses a reflexive view of the position of the researcher with respect to the research.

I am a white woman with a middle-class upbringing, memberships in several environmental organizations, and a graduate degree. These individual characteristics make me demographically similar to many participants in the citizen science projects selected as cases for this study. I take a positive view on the value of citizen science, and science more generally. In particular, I value the scientific endeavor and expect that citizen science projects can make meaningful contributions to both scientific knowledge production and the lives of the individuals involved in these projects. This too is reflective of typical attitudes of citizen scientists (Brossard et al., 2005). I consider these aspects of my position as a researcher beneficial to understanding the participant experience and to some degree inherent in the practice of scientific research. I have taken care not to turn a blind eye, however, to the shortcomings of the cases I study, the assumptions and practices of citizen science, and the scientific establishment more broadly.

In addition to theoretically motivated and practical reasons, my values also influenced the choice of projects, in particular, focusing on environmental science for case study selection. I gave substantially more weight to the research-oriented rationales during selection of cases than my personal preferences, but was pleased that these reasons could also align with personal values. I care deeply about the larger problems that these projects address and

hope the outcomes of my research can support citizen science in tackling important research questions that I am not able to address myself.

Again, my interests and values put me in greater alignment with other participants in these projects, and have helped sustain my enthusiasm for and commitment to the study. Rather than biasing my views toward an entirely favorable position with respect to the cases, I find that my commitment to supporting citizen science has contributed to an ethic of maintaining a balanced perspective. I understand that my criticisms can benefit the development of these projects by pointing out the characteristics and assumptions of these projects, good or bad, that are taken for granted by those closely involved in their operations.

My professional background in nonprofit management supports a deeper understanding of the case study contexts, but also means that I tend to take a positive view of their associated institutions as well. As a former volunteer coordinator, I understand the challenges of managing an unpaid workforce and the potential for mutual value from voluntary work. I believe that these projects have the potential to serve the greater good, like most nonprofit organizations and volunteer-driven communities.

Throughout the study, I took care to observe and consider how my own feelings, assumptions, values, and personal qualities affected the way the research was conducted. Such reflexive practice is an additional method of improving the quality of research. These personal reactions were also recorded in field notes and memos as another form of insight into the context of the phenomenon. I believe that taking my personal characteristics and values into careful consideration with respect to how they affect this study has helped strengthen the research. The disclosure of positionality also provides the reader with information needed to assess the credibility and validity of the work by making apparent the influence of the researcher's perspective.

3.4.3 Going Native

The phrase “going native” is used in anthropology to refer to circumstances in which the researcher becomes a member of the social groups that she studies. Going native was, in fact, a key aspect of the participant observation in this study. To answer the research questions, going native was pragmatically worthwhile and theoretically necessary, so appropriate measures were taken to maintain objectivity and the ability to remain critical.

In order to participate in the projects as a contributor, it was necessary to learn birding, alpine wildflower identification skills, and tend a garden to ensure that my sunflowers would grow. Gardening did not require the level of immersive engagement that would generally be considered going native, and I was already a hiker with good plant identification skills at the start of the research. Reporting bird observations was far more challenging due to the nature of the task, however, and could not be adequately achieved without making a serious effort to become a birder. While going native is considered a detriment to the research quality in some cases, it also provides an unparalleled insider perspective that can substantially enrich the research. For example, in order to truly understand the interests and motivations of birders, and the reasons that the eBird software was designed as it was, it was necessary to become a birder.

From a pragmatic standpoint, becoming deeply involved in the organizer community provided access to the cases and interviewees that simply would not have been available otherwise. The trust developed through relationships with organizers also provided opportunity to obtain frank, honest answers that would have been less forthcoming or entirely absent had the idealized distance between researcher and informant been maintained. When requested to participate in citizen science organizer community workshops and meetings, and to provide feedback on project materials, I felt it appropriate to reciprocate the generous contributions that these individuals had made to my research.

From a theoretical standpoint, pursuing objectivity in qualitative research is paradoxical. The plurality of phenomenal worlds, as discussed by Kuhn (1970), highlights the differences in interpretation that are drawn by members of different groups. Accurately interpreting the positions and nuances of interviewees from the birding community, for example, would have been incomplete without the experiences of birding. Access to this experientially-based knowledge required skepticism of my own assumptions (e.g., that observing birds is “easy”) and pragmatic acceptance of the assumptions of interviewees (e.g., book learning cannot adequately substitute for substantial field experience.) Experiencing the community practices and norms therefore enabled contextually-appropriate interpretation of interview data, some of which would have been naively interpreted without this level of familiarity.

As there was relatively little preunderstanding of the communities of practice compared to insider research in organizational contexts, issues of assumptions and failure to consider alternate framing were less substantial (Brannick & Coghlan, 2007). While there was no intentional effort to influence the organizers’ behaviors at the outset, I found in time that these effects were inevitable as an outcome of seemingly unremarkable interactions. As a person engaging with the citizen science organizer community, but not acting as a project organizer, I had the unusual role of an observer at the community level, which supported the maintenance of an outsider perspective.

The risks to the research of going native are losing objectivity by becoming too close to the subjects and losing the ability to be critical of community practices and perspectives. In response to these concerns, field notes maintained a high level of intentional reflexivity, recording the learning process and changes in perspective as I became a member of these communities. Analytical memos provided additional opportunities to examine the data from a more critical standpoint, particularly through the comparison of data sources. In addition, data collection specifically sought to engage diverse perspectives and opinions, allowing me

access to divergent views of the cases that further prompted critical reflection and reinforced objectivity by presenting conflicting accounts. These practices helped maintain research quality despite the necessity of joining the communities that I studied.

3.5 Summary

This research employed a comparative case study design, for which the conceptual framework developed in Chapter II provided a focus for data collection and initial analysis. Selected concepts from the framework formed the criteria for theoretical sampling to select three cases for in-depth study. The characteristics of these cases were compared to results of a survey of citizen science projects, which verified that the cases represent a cross-section of the larger population while also demonstrating theoretically interesting variations. Field research methods were used to collect several types of data to test and further develop the theoretical framework. As data were collected, analysis began with iterative coding of interview transcripts and description of each case for within-case analysis. Comparisons were drawn throughout the data collection and analysis process, with the within-case analysis completed before cross-case analysis was undertaken in earnest. Combined with the specified research questions, the iterative and concurrent data collection and analysis strategy led to evolving insights on the conceptual framework and emergent themes. The quality of the research was strengthened by several elements of the research design, including data triangulation, creation of an audit trail, multiple stages of participant review, and peer review of findings.

CHAPTER IV

Mountain Watch

4.1 Conservation in the Clouds

In 2004, a plant biologist in the Research department of the Appalachian Mountain Club (AMC) designed a study to monitor the effects of climate change on alpine plants in the Presidential Range of New Hampshire's White Mountains. Phenology refers to the study of the life cycles of organisms and their responses to seasonal changes in their environment. Long-term monitoring of the phenology of flowering plants could help establish how climatic conditions are changing in this fragile alpine ecosystem.

Phenology research has already shown that spring now arrives more than a week earlier at Walden Pond than it did in Thoreau's time (Miller-Rushing & Primack, 2008). The potential impacts of these changes are disturbing. Plants, insects, and animals are ecologically interdependent and their relationships are often time-sensitive. Changes in phenology triggered by climate change could cause mismatches in their life cycles that may subsequently lead to decline and loss of sensitive species.

Plants in alpine environments are of particular concern because of their extremely limited range of occurrence and specialized adaptations to the habitat. Most are very small, compact, and have evolved features that help them withstand high winds and hard winters; even when dwarfed only by climate rather than genetics, alpine plants grow very slowly and recovery

is protracted if they are damaged. For example, krummholz trees of several species (from German, meaning “crooked, bent, twisted wood”) in the White Mountains can take over 100 years to reach the diameter of a human finger, while the same species growing in lowland forests easily achieve this size in just a few years. *Diapensia*, a small pincushion plant that often grows along the edges of narrow high elevation trails and is quickly destroyed by trampling, takes nearly 15 years to grow to the diameter of a U.S. quarter coin and requires an average of 18 years to achieve reproductive maturity.

So how is climate change affecting alpine communities in the Northeastern U.S.? When the biologist plotted out the parameters for his study, he found that it wasn’t possible for one person to gather the data needed to make meaningful conclusions about the phenology of alpine plants. No single person could visit all the necessary alpine monitoring sites, frequently enough and for a long enough period of time—years, in fact—to make the study possible. It would take a lifetime of dedication. No one has the resources for that kind of study.

Not long after, in a brainstorming session with colleagues in the AMC’s Research department, a solution was proposed for the resource gap. The department had been doing some work over the years to look at climate change, including mapping treeline and the distribution of key alpine plant communities. As the Research team discussed other measurements that would be useful to address these question, the concept of phenology monitoring surfaced as an option. AMC staff already knew that visitors to the White Mountains are interested in seeing the alpine wildflowers in bloom; they receive calls and emails from visitors inquiring about the timing of spring flowering every year. There seemed an obvious opportunity to leverage volunteer effort from hikers. Given the scientific research goal of the study, one of the primary methodological questions became: Can hikers in the White Mountains provide scientifically useful data for long-term monitoring? A handful of researchers and educators

devoted portions of their time to finding out.

4.1.1 Project Description

Mountain Watch is a citizen science project operated by the AMC with two sub-projects focusing on air quality and plant phenology. Based primarily at the AMC facilities in New Hampshire's White Mountains, visibility monitoring was the original focus of Mountain Watch. The phenology monitoring protocol is now the primary focus of Mountain Watch, and is also suited for data collection in other forests and mountain ranges in the Northeastern U.S.

The phenology monitoring project piloted its protocols in 2004 and started actively recruiting volunteer participation in 2005. By 2006, hut naturalists provided visitors with interpretive programs featuring the Mountain Watch project in both its incarnations. The Visibility Volunteers use a "view card" to evaluate the clarity of the view from four of the eight huts. Wildflower phenology monitoring requires a more complicated set of activities, however, and although materials are available to let adventurous individuals make self-guided observations, most participants in Mountain Watch encounter the project through an introduction from a hut naturalist.

The Mountain Watch phenology protocol requires participants to find a pre-established monitoring location, identify the target plants for monitoring among others growing nearby, and then identify the stage of plant flowering or fruiting. Participants are also offered the opportunity to participate in locations of their own choice, as long as the monitoring species are present. While a forest flower monitoring protocol is also available to collect comparison data on low-elevation species, the primary focus is on alpine plants, particularly within the Presidential Range of the White Mountains. There were two research questions that the organizers hoped to address with these data: 1) Are mountain plants flowering earlier? and 2) How are environmental parameters related to flowering? These research questions focus

on indicators of climate change through the study of phenology.

The outcomes with respect to overall participation have been encouraging: AMC reported that by 2010, 15,000 hikers had been involved in either contributing data to the project or participating in the related naturalist programs (Buni, 2012). Participation is limited to the growing season, which imposes further constraints given the weather patterns in the White Mountains. In particular, the conditions on Mt. Washington (very close to several monitoring sites) have been called “the worst weather in the world” since 1940 due to its dangerously erratic weather. Wind speeds up to 231 miles per hour have been recorded atop Mt. Washington, and July is the only month with no snowfall. Even the U.S. Forest Service (USFS) signs warn hikers about “the worst weather in America,” noting that numerous deaths of exposure have occurred in this area, even in the summer.

Additional limitations are posed by the target audience, hikers visiting the Whites on vacation, which means that their commitment to the task may be substantially different than participants in the other cases. Given the environmental conditions, the technologies supporting Mountain Watch are primarily paper-based, with a more recently developed subsection of the AMC website permitting online data entry as well. Newer online features map data in real-time, but data retrieval and visualization are still fairly limited.

Mountain Watch has performed impressively rigorous and incremental revision to the participation protocol to support data quality for scientific research purposes by reducing the complexity of the participation tasks. Mountain Watch also serves as a senior member of the emerging phenology citizen science community, offering valuable insights for other groups as they design new projects. Despite these successes, the data produced to date have been of limited research value due to several validity concerns, although these issues are being progressively addressed through ongoing refinements. The project’s organizational support has permitted this slow evolution of the protocol because of the mission alignment with

AMC's focus on education as well as conservation and recreation. Mountain Watch represents a maturing, long-term, place-based citizen science project; it has demonstrated rigorous scientific approaches to protocol refinement and has leveraged organizational resources to expand outreach to a constantly changing participant base.

4.1.2 Organizational Context

The Appalachian Mountain Club was founded in 1876 by Bostonians who shared an interest in mountain exploration (Wivell, 2011). Its organizational mission is to “promote the protection, enjoyment, and understanding of the mountains, forests, waters, and trails of the Appalachian region.” This mission was consistently summarized by interviewees as three focal areas: “our main areas of focus are conservation, recreation, and education. We find that through one of those you’re going to get involved with the other two” (Ledum, 1815–1996). AMC is an institution in its own right; it is the oldest U.S. mountain club, and a recognized leader in its organizational field. As a membership-based organization, AMC has a chapter system similar to other outdoors-focused organizations (e.g., Sierra Club, National Audubon, Adirondack Mountain Club) with 12 chapters that offer local activities and workshops. It is a relatively large organization, with 450 full-time and seasonal staff, 16,000 volunteers, and 100,000 members, supporters, and advocates. In addition to supporting conservation policy and research, staff and volunteers maintain 1,500 miles of trails and work to bring urban and at-risk youth to the outdoors.

AMC has a long and storied history, full of tales of tragedy and triumph. Among its many activities and efforts, the AMC owns and manages several large wilderness areas in the Northeastern U.S., and also operates camps, cabins, and lodges in Massachusetts, Connecticut, New York, New Jersey, and New Hampshire, with the Maine Woods Initiative representing the most recent expansion. Of these facilities, the huts in the Presidential Range of New Hampshire's White Mountains are the most established, notable, and relevant

to the Mountain Watch project.

The White Mountains High Hut system is fashioned after European lodges in the Alps, which provide food and shelter to mountaineers and are often operated by membership-based clubs similar to AMC. The first hut in the White Mountains was built in 1888 between Mt. Adams and Mt. Madison. The huts provide comfortable lodgings and hearty meals to hikers, many of whom would otherwise find the conditions in these locations too harsh and dangerous even in the summer. For example, the Lakes of the Clouds Hut (Figure 4.1a) was built on the shoulder of Mt. Washington in response to the combination of the irresistible lure of the highest peak in the region and resulting tragic deaths of expert outdoorsmen due to exposure. To this day, it has an emergency shelter beneath it that is accessible year-round (described by a tour guide as “nasty” but better than dying.) The lodge and hut system covers multiple public land areas, including the White Mountain National Forest, which attracts around seven million annual visitors.

The White Mountains facilities include two lodges and visitor centers at Crawford Notch and Pinkham Notch. The eight High Huts are spread along a 42-mile section of the 2,200-mile Appalachian Trail and are located approximately six to eight miles apart, from Lonesome Lake near Franconia Notch to Carter Notch Hut between Wildcat Mountain and Carter Dome. They are operated by special permit from the U.S. Forest Service (USFS); the 1999 renewal of the 30-year permit required a multi-year study and environmental impact statement, as the huts are located in environmentally sensitive areas and attract thousands of hikers to the alpine territory above treeline.

The hut lodging capacities range from 36–90 people in bunkhouses or bunkrooms (Figure 4.1d). For a reasonable fee, guests enjoy spectacular views and share multi-course family-style meals in the common areas of the huts (Figure 4.1b). The huts are open for full-service operations from June 1 through mid-September or October, depending on location; three of



(a) Exterior of Lakes of the Clouds Hut.



(b) Dining and common area at Mizpah Spring Hut.



(c) Reception area and kitchen at Carter Notch Hut.



(d) Twelve-person bunkroom at Lakes of the Clouds.

Figure 4.1: AMC hut facilities.

the lower altitude huts offer winter lodgings with reduced service (no meals provided, unheated bunkrooms.) During the full-service season, the huts are operated by a “hut croo”¹ of five to nine caretakers who provide hospitality, day-to-day maintenance, and emergency rescue; they are typically college students or recent graduates. The croos also include a resident naturalist, an individual with expertise or training in natural history, who provides daily educational presentations for guests on topics ranging from moose to how Mt. Washington makes its own weather to Mountain Watch monitoring. A hiker shuttle provides hikers easy access to trailheads without competing for trailhead parking or undertaking joint-jarring road walks along fairly busy highways. This institutional structure and organizational resources—particularly the backcountry facilities with their constant flow of summer visitors—are a substantial asset for the Mountain Watch project.

4.1.3 Data Collection for Mountain Watch

Data collected for the Mountain Watch case study included these sources:

- Interviews
- Participation and observation
- Documents

These data sources complement one another and provide a holistic view of phenology monitoring in Mountain Watch.

Interviews

I conducted five interviews directly related to the Mountain Watch project; each interview was approximately 60–90 minutes in duration. One interview included two interviewees, and longitudinal interviews two years apart were held with one of the organizers. Two interviews were held by telephone or Skype, and the rest were held in person at various AMC facilities and at an ecology conference. Interviewees were selected because they were directly involved

¹The term “croo” is the traditional name for these personnel, and the term is specific to the teams operating the AMC huts.

in organizing and facilitating the project, and the sample represents all but one of the people who were directly involved in project decision-making.

All of the interviewees are (or were) employees of the AMC. Their pseudonyms, based on the Latin names of Mountain Watch alpine monitoring plants, are associated with their roles in Table 4.1. They span the Research, Education, and Programming departments of the organization, which allowed elicitation of perspectives across internal departmental divisions. The researchers provided background on the scientific aspects of the project, and along with the staff person from the Education department, also offered insight into project organizing. The tour guide and former hut naturalist had substantial insight into participant interests and modes of participation, providing a more balanced perspective on organizing and participation.

Pseudonym	Role
Carex	Research scientist
Cornus	Educator
Clintonia	Former research associate & hut naturalist
Geum	Research scientist
Ledum	Tour guide

Table 4.1: Interviewees for Mountain Watch case study

I participated in several events (at least five conferences and meetings) during which informal interviews occurred with three of the interviewees, providing updates to project progress and additional details about the project. Conference attendance also permitted me to interact with the organizers based around the academic presentations of their work in both talks and posters. In addition, I participated in at least 6 conference calls for a related project that involved two of the organizers, who provided accounts to the entire group that reflected on the Mountain Watch project organizing experiences. Finally, the pilot study for the current research included 20 interviews focused on organizers of a phenology network in which Mountain Watch is the most established citizen science project. These meetings

and interviews, while not quoted here, provided substantial background on citizen science phenology monitoring and protocol development, as well as the wider organizational field in which Mountain Watch operates.

Participation and Observation

Participation and observation in Mountain Watch was substantially different from the other cases in this study along several dimensions. The nature of these differences had to do with the unique characteristics of site-based participation in New Hampshire. While participation could have been conducted in other areas such as the Adirondack Mountains, this would have been an atypical form of participation. Participant observation was carefully structured to replicate the most common ways that hikers participate in the project, so the specifics are discussed in detail below.

Participation was immersive, physically demanding, and exhilarating. It required a visit to the White Mountains, located approximately 380 miles from Syracuse, NY. Care was taken to visit during optimal timing for alpine wildflower blooming in mid-June. The huts selected for the visit were also intentionally chosen to provide a sense of the variety and character of these facilities. In addition, an assistant was enlisted for safety reasons: hiking solo is dangerous, and the trails in the Whites cover some of the most difficult terrain in the Northeast. Extensive field notes were made daily at every opportunity throughout the trip.

The weeklong trip was designed to provide experiences representative of multiple modes of participation in Mountain Watch. The trip involved visits to two “frontcountry” facilities, three backcountry facilities, and the summits of Carter Dome, Mt. Pierce, and Mt. Franklin. There were two parts to this visit, one representing the more common independent style of participation, and the second as part of a guided “Lodge-to-Hut Adventure” tour focused on alpine wildflowers, which also incorporated Mountain Watch as a component of the activities.

The independent participation experience began at Joe Dodge Lodge at the Pinkham

Notch Visitor Center with a two-day roundtrip excursion to Carter Notch (elevation 3,288'), located between Wildcat Mountain and Carter Dome and surrounded by two lakes formed by an enormous boulder field. The Lodge-to-Hut adventure was a 3.5-day tour, with seven participants and two guides. Beginning at Highland Visitor Center, the group hiked to Mizpah Spring Hut (elevation 3,800') and then Lakes of the Clouds Hut (elevation 5,012') before returning to Pinkham Notch; it also included a side trip to an alpine bog.

These participation experiences offered the opportunity to directly observe the way that Mountain Watch is presented to hut guests, who are the primary participants in the project. It also allowed direct observation of other participants, both potential participants among the general hut guests, and engaged participants who were part of the tour. As members of the “Flower People”—so dubbed by a hiker who was not part of the group—these individuals were presumably self-selected ideal participants for Mountain Watch based on an interest in alpine wildflowers.

Participant observation in the White Mountains provided an insider perspective that would have been impossible to gain from any other data source. Because there were no avenues for technology-mediated interaction among participants, the in-person experiences were the only opportunity to observe project participants. Both observation and participation revealed a variety of ways that participants can experience both the AMC facilities and Mountain Watch, although the data collection emphasis was on participation rather than observation of other participants. These hikers made an interesting contrast to the participants in the other case studies because they are on vacation, rather than engaging in activities that can be a part of their everyday routines. The experience also demonstrated the consistency and saturation of communication about Mountain Watch to hut guests, which was a truly unique feature of this project. Based on this experience, the nature of another type of participation—independent observations made outside of the Whites—was also clarified,

although not undertaken as part of this study.

Documents

Documents collected for this case consisted of materials provided to participants (primarily guide books, view cards, and data sheets), AMC marketing materials, tutorials, and numerous photographs of the field sites, participants, and location-based promotional materials. Several versions of the participation materials—Mountain Watch guide books and data sheets—were collected because they changed over time. These documents provided reference points for the specifics of participation as well as evidence of the evolution of project communications and participation protocols.

4.2 How Mountain Watch Works

To address phenology-related research questions, long-term monitoring of the same plant species in the same location—sometimes even the same specimens—is required. While scientifically rigorous in comparison to many citizen science participation protocols and therefore more difficult to carry out, the validity of the research is paramount for informed decision-making with respect to land management and policy. This section provides background on the structure of participation tasks and the technologies used by Mountain Watch to support public participation in climate change research. Unlike the following case study description, this section begins with the data contribution tasks rather than the technologies because the contribution tasks affect the technologies more than the reverse.

4.2.1 Data Contribution Tasks

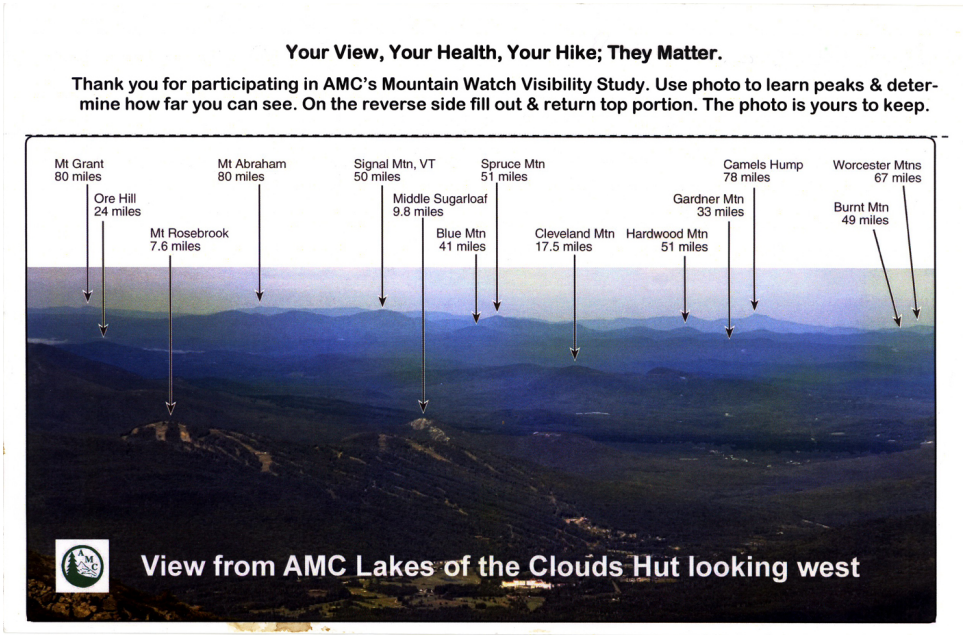
Unlike most citizen science projects, Mountain Watch includes two separate but related research projects: visibility reporting and flower monitoring. While the focus of this case study is the mountain plant phenology portion of the project, the “Visibility Volunteers” monitoring project is also described here for comparison and context.

Visibility Volunteers

The visibility monitoring portion of the project started prior to the phenology monitoring. The purpose of the project is ongoing monitoring of air quality (ozone and particulate haze) through evaluation of relative clarity of the view from high elevations. The protocol originally involved using ozone cards, similar to pH litmus paper, to collect ozone level measurements, and also asked hikers to submit photos from their hikes. This was very popular hands-on activity that concretized an otherwise invisible quality of the environment. It ultimately did not provide adequately scientifically valid data due to the frequency with which hikers encounter inclement conditions that affect measurements on ozone cards. Once the ozone cards were dropped and only photos were requested, participation dropped dramatically.

The monitoring protocol was therefore changed to use a view guide at set locations, which volunteers use as a calibration point for evaluating the air quality conditions (see Figure 4.2, which was the last one remaining at the location from which it was collected.) The data form on the back of the view guide asks for minimal information: date, time, name, whether it is a first visit to the Whites, visual range, presence of clouds, visibility rating, and acceptability of the haze level. The entire process can be easily completed in under a minute. The view cards are distributed at four of the AMC backcountry facilities, and a different view card is required for each location. The tear-off data forms are returned to the same location, which is marked on the form. The in-the-moment nature of participation, its brevity, and the minimal requirements for returning the data form support good follow-through on data form submissions; there is no online data entry for this monitoring project. The image on the front of the card provides a souvenir for participants, reinforcing the value of their participation and providing a reference point for their own vacation photos.

The data form for visibility monitoring includes one opinion-based item, asking hikers whether they believe the visibility indicates acceptable air quality conditions. This data is



(a) Calibration image on the front of the view guide.

Date: _____ **Time of Day:** _____ **AM / PM**

Name: _____ **Are** _____ **Rate** _____ **In your opinion is the** _____

Is this your first visit to the White Mountains? Y / N **Clouds Visible?** **The Visibility (circle one)** **haze level*?**

Visual Range (miles): _____ **Yes** **No** **1 2 3 4 5 6 7** **Acceptable**

(furthest discernible peak) **(1=poor and 7=excellent)** **Unacceptable**

Detach top portion and return to AMC facility



Your View, Your Health, Your Hike; They Matter
www.outdoors.org/mountainwatch

Haze that reduces visibility is caused by small particles suspended in the air that absorb and reflect light. Small particles combine with humidity in the air, further reducing visibility. Air pollution from power plants, cars and trucks, and other man-made sources are the largest contributors to haze pollution.

Many pollutants that reduce visibility can also temporarily impair lung function and aggravate existing health problems. AMC posts the current air quality conditions daily at our backcountry huts. You can also check the air quality forecast at www.airquality.nh.gov.

Your participation contributes to a study that will assess the impacts of air pollution on visibility in the mountains.

The Appalachian Mountain Club is an equal opportunity service provider. The AMC operates Pinkham Notch Visitor's Center and its system of backcountry huts in the White Mountain National Forest under special-use permits from the US Forest Service.



(b) Data submission form on the back of the view guide.

Figure 4.2: The Mountain Watch visibility monitoring view guide.

used to calculate a metric of the visual range that is generally considered acceptable, and has been used in testimony to lawmakers to demonstrate the value that their constituents place on maintaining the Class I airsheds that are protected by the Clean Air Act. The acceptability metric is also the focus of an Environmental Protection Agency revision of National Ambient Air Quality Standards for urban visibility. Collecting thousands of individual responses has meant that even opinion data carry weight for influencing conservation policy.

Although visibility monitoring was the original inspiration for Mountain Watch, it is no longer the primary focus. Plant phenology monitoring, discussed next, has been the central emphasis of the project since 2006.

Plant Phenology

In phenology, the phrase “timing is everything” is particularly meaningful, as the timing of natural life cycles is the focus of this cross-disciplinary research area. The process of participation in the Mountain Watch phenology monitoring protocol has a few variations based on location, but the core task is the same. The hiker picks up a participation kit at one of the AMC facilities—each facility has a location-specific kit—which contains a customized full color field guide to the target monitoring plants and their phenological states, a data sheet, and a pencil in a plastic zip bag (see Figure 4.3). Materials can also be downloaded online before a hike, but do not contain site-specific maps. The data sheet includes a list of monitoring plots with detailed descriptions of the locations for making observations and a topographical map on the reverse side to further clarify the locations. The organizers are not permitted to place markers at these plots due to USFS regulations (Clintonia, 23827–23907), but as will be discussed shortly, observant hikers can use other cues to find some of these locations.

After hiking to one of the specified locations, typically under a half-mile from an AMC facility, participants use the field guide to identify the plants listed for that location on



Figure 4.3: Phenology monitoring participation kits for northern forest flowers (left) and alpine flowers (right).

the data sheet, and then also identify the phenophase/s that those plants display (before flowering, in flowering, after flowering, ripe fruit, and after ripe fruit). There are three sets of six target species for phenology monitoring, one set of five alpine flowers (previously 6; one species was removed due to overwhelming identification challenges) and two sets of six species of forest flowers, with separate lists for both northern and southern species. Hikers can choose at how many and which of the specified locations to make observations, and can also report on these target species for optional locations of their own choosing.



Figure 4.4: Mountain Watch data sheet drop box at Pinkham Notch Visitor Center, located at an information desk.

Using the provided pencil, the participant fills in the data sheet for the specific location and then tucks it away until reaching their next observation point or AMC facility. When the participants have finished making observations, they can return the completed data sheet at any of the AMC facilities, where clearly marked data sheet drop boxes are positioned in

high visibility locations (see Figure 4.4). If this is inconvenient, participants can also submit observations online through the Mountain Watch section of the AMC website.

4.2.2 Meta-Contribution Tasks

Meta-contribution opportunities, where participants can make contributions by supporting the efforts of other participants, are extremely limited. In fact, interviewees made no mention of engaging volunteers in any aspect of the project other than data collection. The primary reason for this, despite resource limitations and competing demands on staff time, seems to be that more relatively low cost internal labor is available to Mountain Watch organizers compared to most other projects. These human resources are primarily hut naturalists and interns, all of whom have more time to devote to learning how to do more complex and interdependent tasks than would be expected of volunteers.

4.2.3 Mountain Watch Technologies

To support the scientific interests of the project organizers and the needs of participants, several technologies are used in Mountain Watch. Mountain Watch contrasts sharply with the other cases in the degree to which it relies upon paper data sheets, which are an important technology for many field-based monitoring projects. With very few exceptions, paper is the only practical approach for field-based data collection at this time, as will be seen in the other cases. However, for most Mountain Watch contributors participation involves *only* paper data sheets and very few individuals submit data online, while in the other cases data submissions are exclusively or predominantly online. Climatic conditions mean that the packaging of the data sheets in a plastic zip bag, while not only practical for inclusion of the field guide and pencil (which cannot otherwise be assumed to be at hand), also protects the paper from the inclement weather that is a nearly inevitable feature of visits to the Whites.

Paper data sheets and field guides are the most important tools for Mountain Watch

participation. The reason is simple: bad weather is bad for electronics, and inclement conditions are the norm rather than the exception in the Whites. Much of the evolution of the project's participation protocols, described later, is reflected in the data sheets. There are multiple types of data sheets currently in use, which represent progressive refinements. Custom field guides accompany the data sheets (shown in Figure 4.3) because most individuals are unfamiliar with not only the monitoring plants, but also their phenophases. The field guides, developed in collaboration with AMC's Education department, contain color photos and extensive details about how to make these identifications, which can be quite nuanced. For example, as bunchberry develops, it has white outer leaves that appear to a casual observer to be petals and therefore suggest that the plant is in bloom, but the plant's true blooms appear later, a tiny cluster of flowers at the center of these false petals.

The paper data sheets (see Figure 4.5) take several forms. Most are specific to the location from which they are distributed so as to include maps and descriptions of the nearest monitoring plot locations as well as the species that can be found there. There is also a general version of the form without locations marked, and extra space for describing monitoring locations. This is the version of the data sheet that predated the location-specific data sheets, and is available at the AMC visitor centers and online for use in locations both within and outside of the White Mountains. The online version of the data sheet does not have a map of pre-specified plots, and volunteers were encouraged to draw or provide a map for data submissions. As previously mentioned, the data sheets can be returned to AMC facilities, which substantially improves the return rate because participants do not have to go out of their way to submit data; these data are entered into the database by interns.

Online data entry is a relatively new feature of the Mountain Watch webpages, which are integrated into the larger AMC website, maintaining the organizational brand. This technology is not required to participate and was added after the project launched, unlike

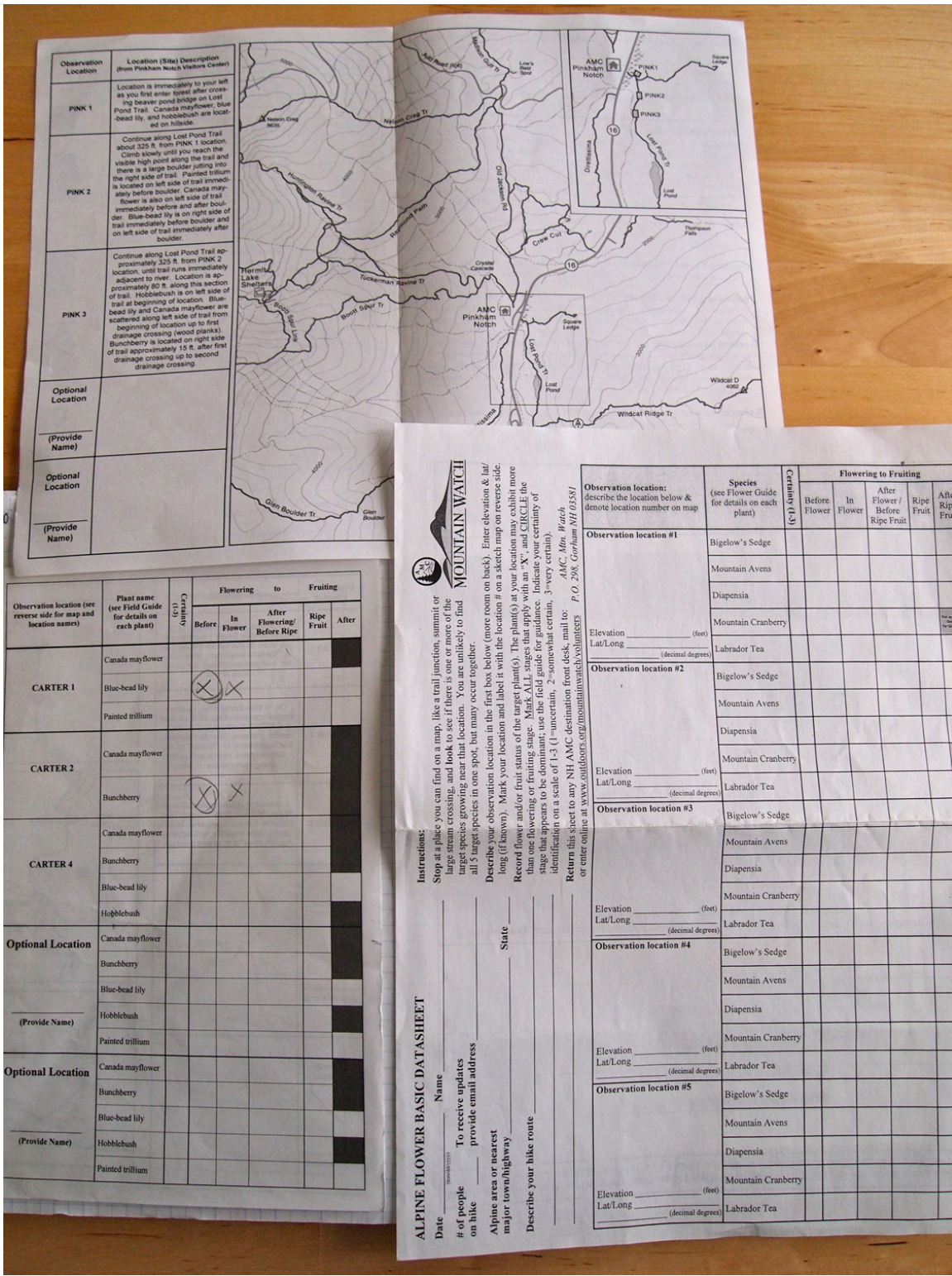


Figure 4.5: Mountain Watch data sheets with plot descriptions and map (top), a location-specific form (left), and generic form for alpine locations (right).

the other cases in this study which are fundamentally dependent on web-based systems and required online data entry functionality from the very start. The development of online data entry has been incremental, but permits participation from a wider range of forest regions and mountain ranges outside of the Whites (e.g., the Catskill, Adirondack, and Green Mountains). This broadens the potential contributor base and provides the capacity to substantially increase the scale of participation. The online data entry form for Mountain Watch observations (see Figure 4.6a), however, does not exactly match the paper data sheets; it is logical in terms of the data entry itself, but the implications of variations between paper and online data submission are unknown.

Species	Certainty of ID	Phenology Stages					
		Before Flower	In Flower	After Flower / Before Ripe Fruit	Ripe Fruit	After Ripe Fruit	
Bigelow's sedge	<input type="text"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
dominant:		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
Mountain avens	<input type="text"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
dominant:		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
Diapensia	<input type="text"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
dominant:		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
Mountain cranberry	<input type="text"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
dominant:		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
Labrador tea	<input type="text"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
dominant:		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
Alpine bilberry	<input type="text"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
dominant:		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	

Enter additional observation from same hike

(click once)

(a) The online data submission form for alpine plants.

Program	Protocol*	Adopt-A-Peak	Mail Kit
Alpine Flower Watch	Basic	<input type="checkbox"/> Yes, location:*	<input type="checkbox"/>
Forest Flower Watch	Basic	<input type="checkbox"/> Yes, location:*	<input type="checkbox"/>
Visibility Volunteers	Non-Participant	<input type="checkbox"/> Yes, location:*	info
Fall Foliage	Non-Participant	<input type="checkbox"/> Yes, location:*	<input type="checkbox"/>

Relevant Experience

Better than average knowledge of forest flower species. Participated in Alpine Wildflower Lodge-to-hut Adventure in June 2011. Learned to identify 63 species from New England Alpine Summits field guide (excluding mosses, liverworts, & lichens.)

Previously submitted experience:
I know a handful of common forest wildflowers by sight and keep my field guide in the car for easy reference. I'm good at noting relevant field marks when I'm not certain of a species.

Password

Current Password:

If you would like change your password, please enter the new password below:

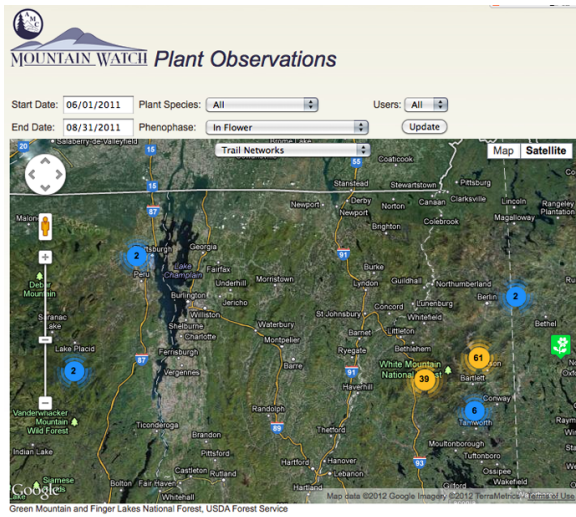
(b) The online account profile form.

Figure 4.6: The Mountain Watch online participation forms.

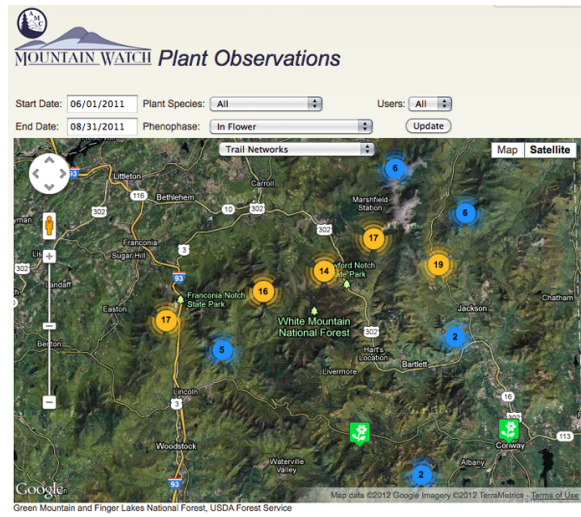
The online data submission option has surfaced a notable difference in participant skills: those who participate independently seem to put more effort into providing quality data based on the completeness of the records and level of detail they provide with the observations. In the online account creation form, contributors are asked to describe their botanical background, as a few participants are in fact professional botanists (Carex) (see Figure 4.6b). Given the opportunity, the researchers would pay more attention to these individuals as a method of filtering data for quality, but there is no such detail available for most participants. As other citizen science projects (including eBird) reported, asking contributors who use the paper data sheets to provide more data in the moment would likely result in declining rates of participation or more incomplete data. Turning self-described experience into usable data for profiling contributors according to skills and prior experience was considered too difficult. Currently, organizers would have to manually classify individuals, but adding a profile categorization option to the data sheet has been considered.

In addition to online data entry, organizers added a real-time map of phenology reports in 2011. The public-facing map shows only recent “in flower” observations, but there are numerous variables that logged-in users can adjust to view data (several such views are shown in Figure 4.7), including viewing only their own data. The website also shows an animated image of a *diapensia* plant moving through its phenophases. These features are the extent of data access and interaction for website users, but the map is more advanced and interactive than those available for most other citizen science projects, typically static arrays of dots on a large-scale non-resizable map. The map view may be an adequate mode of data access for many participants.

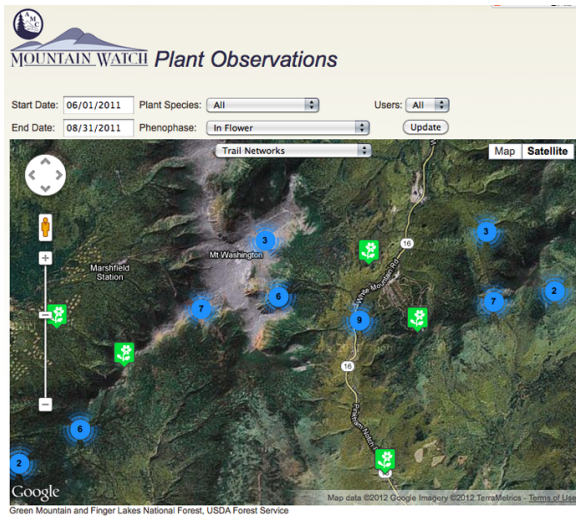
Despite the investment to produce a quality data visualization for participants—important in its own right—the organizers reported that there is no convenient reporting; data can only be retrieved through direct database queries and manipulated with other software. This was



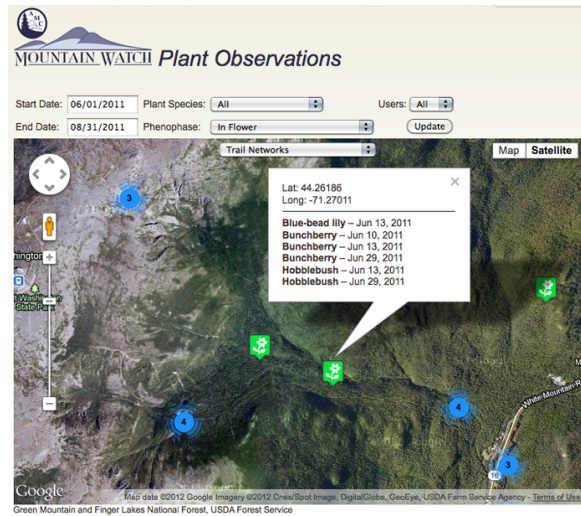
(a) Northeast regional observation map.



(b) White Mountains observation map.



(c) Mt. Washington observation map.



(d) Site-level observation map.

Figure 4.7: The Mountain Watch online interactive maps at several geographic resolution levels.

likely due to a combination of factors, not the least of which is the ongoing refinement of the participation protocol and variable data quality to date, which interfered in prioritizing development of data reporting features.

In addition to these primary technologies, AMC researchers deployed “PhenoCams” in 2009 to collect data remotely; also known as plantcams, these digital cameras are made for ongoing monitoring in outdoors environments (see Figure 4.8a), and grant funding permitted expansion of the project in 2010. This technology investment was spurred by three factors: rapidly decreasing costs and wide availability of remote camera technologies; potential for verification of participant data reliability; and more importantly, the ability to do “off-season” monitoring. Some of the phenological events of interest occur before the huts are open for the full-service season (starting June 1) and weather conditions are typically so severe through late spring that there is simply no one around to monitor the plants. Most of the plantcams were located at established monitoring sites in 2011, so they also have the potential to serve as a point of verification for citizen science observations that may appear questionable. An additional benefit of the automatic data collection instruments shown in Figure 4.8 are that they can be used by volunteers to identify monitoring plots that cannot otherwise be marked, as they are located near trails.

An interesting application of the plantcam images has been the production of time-lapse videos that show the progression of plant phenology for alpine and forest flower species. One of these videos is a simple, brief demonstration of phenophase changes in *diapensia*, a particularly popular alpine flowering plant (<http://www.youtube.com/watch?v=8UrOfkOufPE>). Another type of video, however, is targeted toward volunteer training and includes descriptions of the phenophases for identification purposes. For example, bunchberry can be difficult for untrained volunteers to correctly identify flowering phases due to its false petals (<http://www.youtube.com/watch?v=S81jPngzqog>). These videos demonstrate the use of

technologies that were initially deployed to extend the time range and verifiability of data collection being used for outreach and education as well.



(a) A plantcam at a monitoring plot.



(b) A HOBO[®] data logger for air and soil data.

Figure 4.8: Automatic data collection instruments.

On the other hand, there are several challenges to using the photos, as they require human analysis and the cameras can only be focused on a small patch of plants. AMC is working with a group of researchers to explore the possibility of using a web-based game as an approach to analyzing these photographs. This would add a new dimension to the citizen science project, as it could serve as a type of online training for individuals who participate in activities like organized alpine wildflower hikes or repeated monitoring, as well as offering a solution for ongoing data processing in the future. It would also expand participation opportunities beyond current geographic constraints and could engage a new audience, extending the broader impact of the project. This partnership is one example of the types of organizational and institutional arrangements discussed in the next section.

4.3 Organizing in Context

The organizational context in which Mountain Watch was developed has significantly impacted the project's evolution. Access to organizational resources such as those previ-

ously described has helped maintain the project, despite a lack of consistent funding to ensure sustainability. This form of organizational support substantially increased the reach of project communications. In the broader organizational field, Mountain Watch organizers coordinate with other mountain clubs in the region as well as a phenology monitoring network, though these collaborative efforts are also limited by constraints on project leaders' capacity. This section delves into themes associated with organizing the project, particularly organizational embeddedness and the relationship of Mountain Watch to other members of its organizational field.

4.3.1 Organizational Embeddedness

The concept of embeddedness typically refers to the degree to which individuals or firms are entwined in a social network or organizational field (Granovetter, 1985; Putnam, 2001). In this case, the term refers to the degree to which a project is entrenched in its organizational or institutional context. The level of organizational embeddedness observed in Mountain Watch was unique to this case.

Mountain Watch is the product of direct collaboration between the Education and Research departments at the AMC. As an organizer explained, “we all got together and sat down and talked about, how can we bring the work the research department does more into our education outreach to our members? ... We decided it had to be something that really is based in our science” (Geum 2011, 1107–1434). This is a good example of incorporating education and outreach into a scientifically-focused project, which is a known best practice for citizen science projects (Bonney et al., 2009).

Given the small size of the Research department in the organizational context (five individuals), the internal partnership has also extended the organizing skills that have been brought to bear. This has been an advantageous arrangement, as it has influenced the development of well designed materials that support participation and training that hut croo

members received. The dual focus on science and education has also contributed to the project's sustainability from an organizational standpoint, which is discussed further later.

Although the project started off with the support of multiple departments within AMC, organizers reported that it has taken a long time to take hold within the organization. This was related to institutional commitment to supporting the project, the strong existing cultural component of hut crews, and AMC's traditions as a volunteer organization. "It took us some years to get over certain institutional hurdles...it's actually not easy. But we have definitely gotten over that hurdle in our [New Hampshire] huts. Have we gotten over that hurdle in others, like Maine Woods? No." (Geum 2011, 12226–12671). The slow progression of integration of the project into broader organizational activities and culture has been effective in the White Mountain facilities (as discussed later), but surprisingly, involving AMC's wider volunteer base has been less successful: "AMC is a volunteer organization, that's what part of our mission is, that's part of...how we've survived. But even to try to tap into those volunteer resources has been really difficult internally" (Geum 2011, 16847–17081). This situation is an outcome of existing traditions within the AMC, as well as competition for volunteers as resources, an issue identified in other citizen science projects that are place-based and embedded within institutions that have strong volunteer management norms, e.g., the National Park Service.

Although mission and organizational embeddedness provided advantages to Mountain Watch, they also posed some constraints related to participation and scientific interests. "I see it [the project] as restrained by our mission and our audience and that we are not trying to do citizen science for citizen science. We're trying to do something that is helpful to our science department and our mission" (Geum 2011, 7072–7337). This quote identifies a specific challenge related to the project's ability to demonstrate the scientific outcomes that would justify dedicated funding for the project. Achieving these outcomes have been

complicated by the limited target audience: “We have hikers, we think that they can help us. ... Of course they are our audience, I mean that’s who we [AMC] are. We can’t change that. We can’t change who our audience is” (Geum 2011, 6601–6830).

The slow pace of integration and development of organizational support for Mountain Watch and other citizen science projects in similar contexts seems to demonstrate a reluctance to make an institutional commitment to long-term projects which have yet to demonstrate strong scientific outcomes. Results often require more time to generate in monitorings project than some other types. Other aspects of organizational embeddedness were the slow but successful integration of the project into the hut croo culture, and the lack of dedicated resources for project sustainability.

Hut Croo Culture

The culture of AMC hut croos could make an interesting study in itself. This program for staffing the huts has been developed over 50 years, involves rigorous screening of candidates, and has a number of strong traditions. These traditions include, for example, nighttime raids to steal symbolic treasures from other huts, after-dinner skits and songs to inform and entertain hut guests, and specialized slang for aspects of hut croo life (over 40 terms are listed in a glossary of “Mount Vernacular” (Wivell, 2011).)

The traditions in this singular organizational subculture are slow to change: one organizer referred to the introduction of hut naturalists approximately ten years ago as a “new” innovation. A project leader also commented on the lengthy process of integrating Mountain Watch into hut croo activities: “it’s taken a couple of years, because it just takes a little while to get the information to take hold in the huts. And then once it is, they’re good about sort of, ‘oh yeah, we did that last year,’ and they’ll repeat it” (Cornus, 23083–23319).

With respect to the integration of Mountain Watch into the hut croo responsibilities, the collaborative relationship between the Research and Education departments led to the

addition of more intensive training for hut naturalists. All hut croo members receive some limited natural history training so that they can answer basic questions. The hut naturalists serve a special role with Mountain Watch by making regular phenology observations at permanent plots near the huts and by incorporating Mountain Watch into their presentations to hut guests. The importance of hut naturalists to the entrenchment of Mountain Watch in hut activities was also linked to the current lack of Mountain Watch adoption in AMC's more recently established Maine Woods facilities, which have a smaller all-purpose croo and fewer interpretive staff to directly communicate with participants about the project.

The role of hut naturalists makes a good example of successful efforts to embed Mountain Watch into organizational processes and culture. The details of training for naturalists was described by an organizer:

We have eight naturalists, one at each hut, and they go through what's called our Gala Training in the spring. So at that time, Carex, who is in the Research department, comes up and talks to all of the naturalists, and shows them what's expected of them in terms of monitoring plots, shows them where the plots are. ... And then we have sort of a sample example of people giving a presentation of...the way we introduce it to the guests that are visiting, so they have chance to kind of see that in action. And then following that up, all of the hut croos go to their individual sites, and Research again will follow up with a site visit to each of those folks, to say 'here are your sites.' Then I follow up with 'okay, where are your packets, how are you introducing this to different people?' I watch to see that they're doing that, give them feedback. (Cornus, 21770–22764)

This mode of incorporating Mountain Watch into hut croo training supports communication and outreach that support ongoing participation by a constantly changing audience, as well as data validation through their own data collection activities. Croo members also mentioned the incentive of an ice cream party, a particularly prized reward due to the rarity of frozen treats in the backcountry, for the croo that brings in the most Mountain Watch observations.

As the project evolved, additional support for the hut naturalists' Mountain Watch out-

reach has seen further development.

One of the things that we've learned is making sure they [naturalists] have a story to tell...it's less of a challenge for them. ... So if they wanted to focus in on what Native Americans did with bluebead lily or whatever, they have...resources to build the naturalist and education stories. ... You have to tell the story if you're going to engage the general public in this process. It's not only good for the naturalists who are promoting the program, but it makes it interesting for the guests who volunteer. (Geum 2011, 24340–25168)

This example shows the ongoing development of the project and the organizers' commitment to promoting communication, awareness, and the personal interests of participants. Despite the dedication of the staff who have developed and nurtured the project, the project still suffers from similar resource constraints observed in other citizen science projects.

Dedicated Resources for Dedicated Staff

The primary resource constraint that prevents further development of Mountain Watch is organizer capacity. Project leaders openly admitted to underestimating the effort required to organize a citizen science project: “It's a beautiful concept...but when it comes down to implementation, I mean it's just so much bigger than you would ever imagine when you first come up with this great idea” (Geum 2011, 18395–19302). This frank statement on the demands of organizing volunteer participation in scientific research is reflective of the reported experiences of other citizen science project leaders more broadly.

The main organizational resource that the project draws upon is low-cost or sunk cost labor. These human resources include the hut naturalists; a backcountry education assistant who supervises naturalist programming and helps crew members work on their presentations; and interns who enter data from paper data sheets into a database, process plantcam images, and assemble kits of participation materials.

Financial resources that would support dedicated staff time, as opposed to the fragments of time squeezed from individual staff schedules, present the biggest limitation on project

growth.

In an ideal world, we could have a full-time person to do volunteer management, as we do for our trail program. Unfortunately, we haven't had the funding to do that. So while we've learned some lessons about what works, we're still fairly understaffed...and that's the consistent problem...with these types of programs. Even if you get an initial grant that provides some funds, how do you keep things going as that grant goes away? ... Our facilities provide us the benefit of...a trained naturalist on staff that's doing the monitoring and they can lead the guests to the actual monitoring site and walk them through the materials, so they [hut guests] are actually getting some direct initial training to gain confidence to go out and do it on their own. (Geum 2009, 14653–15815)

The role of the hut naturalists, as previously mentioned, is a significant asset in addressing the shortfalls in volunteer management with respect to communication and training.

The natural history presentations of Mountain Watch to hut guests represents an intermediate approach between the dominant participation training strategies in citizen science, which are usually entirely absent, minimal and self-guided, or else detailed and intensive. By comparison to the other cases in this study and to most technology-mediated citizen science projects, the training for Mountain Watch participants is substantially higher for contributors who are introduced to the project through hut naturalist presentations. Compared to more established volunteer monitoring projects, however, the training for Mountain Watch participants is very limited. The former research associate observed that “one of the things about Mountain Watch is that of all the citizen science projects that I was researching for my Masters, Mountain Watch had the least amount of training for their volunteers, I think by far” (Clintonia, 51475–51957). More traditional volunteer management for long-term monitoring would involve careful training for volunteers and work to retain a committed core of contributors, but this was not a realistic expectation for Mountain Watch: “We are locked into our audience, and...I have no resources to hold a two-day training, or even recruit those [local] volunteers to get them to that training” (Geum 2011, 16037–16391).

Although resource constraints were an issue for Mountain Watch, as with most citizen science projects, the issue of long-term sustainability was rarely raised by interviewees. This is likely because the project has additional value beyond the scientific interests, and the educational value of the project also dovetails with organizational mission. In this respect, Mountain Watch's organizational context is similar to eBird. The match of project activities and potential outcomes to mission meant that in both cases, some degree of organizational commitment to the concept kept the project going through difficult initial years. Mountain Watch was organized by the Research department, so AMC's needs for a scientific basis for conservation linked the project even more closely to the organizational mission. In addition, the project goals also appear to be well aligned with the personal interests of the individuals and communities that participate in the project. Despite the apparent alignment and potential appeal, the match of hikers' skills and interests to the participation activities have implications for scientific outcomes, a topic discussed in later sections.

4.3.2 Leadership in the Organizational Field

Mountain Watch has demonstrated leadership in its organizational field through partnerships with citizen science phenology monitoring projects, which are a fairly recent development in the U.S. The project's interactions with its organizational field have been focused in two areas: enlisting the support of other mountain and trail organizations, and participating in the development of national and regional phenology monitoring networks.

Mountain Watch organizers have worked with other membership- and volunteer-based mountain clubs and trail groups to expand Mountain Watch monitoring throughout the Northeast region. In particular, the project leaders were successful in enlisting contributors from the Green Mountain Club, Baxter State Park (the second-largest alpine area in the region), and members of the Adirondack Mountain Club (ADK) through their Alpine Steward program. The ADK, which incidentally shares the same mission focus on conservation, edu-

cation, and recreation as the AMC, was involved in scientifically-rigorous monitoring, which was seen as an entry point to broader outreach to and adoption by the ADK membership:

Adirondack Mountain Club is doing permanent plots versus more general plots. ... They're contributing as a club to sort of the real high-end data set with permanent plots, which is great. The next step is to get them to utilize the materials to promote with their own outreach to other hikers. (Geum 2009, 21263–21790)

These relationships provided opportunity to expand monitoring to other alpine zones outside of the White Mountains, as well as the forest flower monitoring, to collect comparison data for climate change effects at different elevations and locations.

The Mountain Watch organizers' participation in phenology monitoring networks has been more altruistic than self-serving. They receive relatively little direct benefit, but have provided advice and guidance based on their years of experience in running a location-based phenology monitoring project. They have also worked toward sharing data with the USA National Phenology Network (USA-NPN) and the Northeast Regional Phenology Network. This posed some challenges when it came to integrating AMC's existing phenology monitoring program into a newly-formed network due to the variations between monitoring protocols, which has a substantial impact on the comparability and long-term interoperability of the data. Nonetheless, the project organizers have remained involved in the development of these networks out of commitment to supporting phenology research on a larger scale.

The primary impact of these partnerships on Mountain Watch has been ensuring that the phenophases being monitored in each network are suitably aligned:

We've made sure that...our phenophases are integrated enough that we can share information. ... What I'm unsure about is how integrated our two programs will become in the future. ... Where we can, we've coordinated on protocols and so we can still have that data exchange. (Geum 2009, 7717–8658)

Full integration of Mountain Watch with other projects in the broader network-based efforts seems unlikely, particularly because AMC's project has addressed a need for monitoring in

areas that are not prioritized by other groups. The project organizers saw Mountain Watch playing a valuable role with respect to the USA-NPN and other organizations partnering in phenology networks: “we’re still filling a niche that they are not going to get to, which is trailside and mountains” (Geum 2011, 19514–19657). This potential for an unmatched contribution to regional phenology monitoring efforts has been another argument in favor of continued organizational support for the project despite current uncertainty over the scientific usefulness of volunteer-generated data.

An additional outcome of integrating Mountain Watch with the developing regional phenology network was to establish enough uniformity not only for data exchange, but also for developing an identity and familiarity for participants:

I think it would be useful to have some identity to it, so that people understood that these are all partners in one big effort. ... If they [AMC members] go to a National Park and they participate there and they see AMC’s a partner, and then they come to an AMC facility and it’s similar enough that it’s not a completely different experience. ... I would like to see enough consistency so that people see the connection between the programs. (Geum 2009, 26898–28453)

The desire to develop a stronger interorganizational identity around citizen science phenology monitoring indicated an awareness on the part of organizers that hikers seek out special places that cross many boundaries, both geographical and organizational. Creating partnerships across these boundaries is perceived to have potential for increased participation, and the additional prospect of increased data quality based on the common assumption is that repeated participation will increase skills and thereby data quality. The next section discusses the implications of these organizational characteristics for the scientific work of Mountain Watch.

4.4 Climate Science for Conservation

The primary goal of Mountain Watch has always been to generate data for scientific research that can inform AMC’s conservation efforts. The citizen science approach was taken because it was not feasible for individual researchers to accomplish the data collection required for this type of research. As a result, another distinguishing feature of Mountain Watch is explicit integration with conventional scientific research.

The AMC research staff described their efforts to monitor climate change in the alpine region as a multi-faceted approach; the phenology citizen science project is only one component of the larger effort. Soil and air temperature monitoring have permitted the scientists to develop models for alpine species to better understand how factors like temperature, elevation, and latitude affect the plants. The phenology project fits into this research by complementing the automatically logged air and soil data with observations of the changes to nearby plants. The scientists are also taking multiple approaches to collecting phenology data. The focus of these interconnected research projects is on alpine plant conservation, guided by AMC’s purpose and shaped by the characteristics of the mountains.

This section discusses several aspects of the scientific research, particularly the organizers’ rigorous research-based approach to revising the participation protocol, and the multiple layers of data collection employed to address a variety of research challenges.

4.4.1 Science and Participation: Rigor and Revisions

For many citizen science projects, initial scientific and participation design decisions often need to be changed after the project launches. As a Mountain Watch organizer reflected, “you almost really have to be a scientist in citizen science to understand what you’re getting into” (Geum 2011, 13602–13704). This sometimes lengthy revision process was clearly an aspect of organizing a citizen science project that these organizers—and many of their peers—did

not anticipate.

Mountain Watch has implemented multiple revisions since it was initiated. The primary concern that these changes have addressed was the ability of hikers to follow the participation process well enough to generate data of adequate quality for the scientific interests. Although there is no dedicated ongoing funding, the organizational commitment to the scientific goals and educational value of citizen science have allowed the organizers to progressively address concerns about data quality. According to an organizer:

There are still a lot of issues with volunteer data, but I think that AMC is kinda committed to working through that, and figuring out what they can do to make their data better. They're not just willing to settle on this trope that the more data we have the better, it doesn't matter what the quality is, if we have like enough data, it will even itself out. (Clintonia, 45191–45644)

This is essentially the opposite position from the approach that eBird has taken by focusing on the power of large-scale data to overcome error. The very different geographic scale and audiences of these two projects suggest that it is not realistic to expect that Mountain Watch could achieve the same type of results with respect to the volume of data, which justifies a focus on improving data quality over quantity, a topic that will be discussed later in this section.

Data quality is particularly important to Mountain Watch because it is required to address the project's explicit research questions, the investigation of which preceded the citizen science project component. Creating a participation protocol that can achieve these goals required several adjustments:

We would like to remain science driven, we have a lot of education staff that have helped us dumb down some of the stuff. But you know, when we first came at this as scientists, we definitely had a different approach, and they helped us move it more towards something that is more palatable for a citizen. (Geum 2011, 14773–15213)

The partnership between the Research and Education departments has therefore supported

not only communication, participation, and organizing, but also provided non-researcher insights into how to modify conventional monitoring research designs to make them suitable for members of the public.

For example, the organizers eventually realized that even the phenophases required simplification:

We've had to modify our broader citizen data collection, that's where we came up with before flowering, flowering, and after flowering, because you know, we went through a couple iterations, and then that was as much as ...they could handle that we could get good data out of. So ...we kinda came around to the fact that if we're going to engage that broader, less botanically-knowledgeable membership of ours that we really want to engage, it has to be simplified. (Geum 2009, 22158–22877)

In addition to changes that simplified the phenophases into more easily observed stages, later modifications to the protocol included expansion of the phenophases to cover the fruiting stages of development as well as flowering. The duration of the hiking season is longer than the flowering stages, so this change permitted participation by a larger number of individuals, while also providing additional research data.

As with many monitoring projects, the process of refining the protocol has taken substantially longer than organizers had foreseen: “I feel like we're getting close to a sort of finalizing our protocols, which you wouldn't think it would take five years” (Geum 2011, 4980–5119). The time required to make these revisions was considered a function of the resources available for organizing the project. Mountain Watch is a multifaceted research project incorporating conventional science and citizen science, and the funding for coordination and materials to support the citizen science portion is most limited.

This is...why our program has taken so long to evolve, is because we never had from the get go, some real significant resources, at least at the time when we learned so much. You know, it's like we had an initial \$50,000 [for the entire project]...it doesn't even seem like it was that much. It did not go far and it was in the beginning before we really knew a lot. So I would've spent that money completely differently, had I known what I know now. (Geum 2011, 17149–17981)

This quote not only makes a point about the influence of resources on scientific outcomes, but also demonstrates that the knowledge resources to design a successful project were lacking, a complaint also raised by the Great Sunflower Project. Mountain Watch similarly shows that developing a successful project requires a more substantial initial commitment in terms of startup funding than is available to most citizen science projects.

Despite these constraints, the Mountain Watch organizers persevered, driven by their commitment to doing research that was considered important both scientifically and with respect to mission, but that also required involving non-scientists. With a scientific research mindset, the project organizers set about determining the changes needed based on rigorous scientific research, complete with statistical analyses that demonstrated conclusively that there were several problems related to task complexity. The participation process, while relatively straightforward in description, involves a number of steps and relies on strong observational skills. The task of finding a location, which one might presume to be a natural skill for hikers who navigate wilderness trails, is remarkably challenging when visual cues are limited or spatial judgment is required. In addition, identification tasks of any type are a known data quality issue for many citizen science projects, as reliability is often questioned when most participants cannot be expected to have a strong background in species identification (birders are a notable exception.) These specific concerns are also addressed in further detail later in this section.

Developing a viable protocol required substantial human resource investment, however, including a full season for a research associate to investigate data quality. While AMC is fortunate to have a good supply of individuals capable of this type of work, justifying and acquiring funding to make such a rigorous study of the participation protocols and their impacts on scientific outcomes was a challenge. Nonetheless, this approach to project development was representative of a substantial commitment to the long-term goals of the project.

The remainder of this section addresses the organizers' choice to emphasize data quality over quantity, issues related to geographic and botanical knowledge that were identified as key issues with participant data quality, and the resulting changes to the data sheets.

Quality Over Quantity

The organizers' decisions to focus on improving data quality before addressing quantity were motivated by multiple considerations. The question of whether a large enough data set can offset error has not yet been answered:

Is the error so big that it's giving you misinformation, or is it the whole principle of...crowdsourcing a lot of information, will it [error] be reduced enough so that you can still make sense of it? ... I think we'll be going at it as, how well did they do, when we told them where they [the monitoring plots] were, we told them which plants they'll see? (Geum 2011, 38617–39058)

Given this uncertainty, refining procedures for optimum quality was a sensible first step. Over time, the Mountain Watch organizers have pinpointed the issues of geographic and botanical knowledge (discussed shortly) as their primary concerns with respect to data quality, and have developed strategies to evaluate the effectiveness of changes to the monitoring protocol.

Temporality also plays a role in these approaches to improving data quality. In general, the design of phenology research requires longitudinal data, which means that spending time perfecting the protocol is a worthwhile investment in the long run: “all of our sort of missteps and our learning processes along the way are valuable in the sense that we should ultimately be able to get really high quality data from most of the participants” (Ledum, 20911–21759). The seasonal nature of participation and the environmental conditions in the White Mountains also influence data quality, as participation during the early stages of plant development are limited. These challenges were mitigated by the addition of automated monitoring in 2010, a later topic of discussion.

When it comes to research results and scientific outcomes, the work generated by Mountain Watch researchers for presentation at scholarly meetings has focused on the design and refinement of the participation processes. This is a direct result of constraints on project organizers and also represents a valuable contribution, but the intended scientific results have been limited. Nonetheless, preliminary results from the larger research project into which Mountain Watch is integrated have been interesting:

Our research has shown that the top of Mount Washington, the alpine areas in the Northeast, actually aren't seeing the same strength in warming, or the same rate in warming that the region is, and so...it's even more important to get the lower elevation observations. (Geum 2011, 11233–11572)

These observations have motivated further development of forest flower monitoring protocols to complement the more established alpine phenology research. As yet, resource constraints have prevented further outreach to develop more participation in these lower elevation protocols.

Like other citizen science projects, an economic argument has to be made to sustain Mountain Watch as a research project. One Mountain Watch organizer explained the focus on data quality in terms of return on investment: “if we're giving out thousands [of kits] a season, those costs do add up. ... That's another reason to make sure the data that we're going to get is viable, so that...we're getting our money's worth” (Ledum, 28122–28388). This sentiment has been echoed by organizers of other projects set in institutional environments where a bottom-line judgment on whether to continue supporting a citizen science project is dependent upon the scientific value of the data generated by participants. The question of whether resources going into a project are paying off poses pernicious challenges to projects for which scientific outcomes are the primary goal. This seems to be particularly true for seasonal projects, which may require substantially more time to craft a usable protocol. For Mountain Watch, the two major complications for improving data quality have been

participants' botanical knowledge (or rather, lack thereof) and their skills in identifying and describing geographic locations, discussed next.

Participant Skills: Location Identification

Compared to the other cases in this study, higher precision of locations for monitoring was of greater importance for Mountain Watch because of the research requirements for repeated observations and geographic covariates such as elevation, aspect, and slope. Prior to 2010, monitoring was permitted anywhere along trails in the Whites, and participants were provided with space on the data sheet to describe their location. The degree to which hikers were able to provide a precise description of their location was widely variable:

We had three categories of geography. There is the best guess category, which was the really bad descriptions...the general locations, so that was a little bit more precise, and then the precise locations. So the precise locations were always at trail junctions and summits. There just weren't any descriptions that were precise enough to find on the map again without those kinds of landmarks. (Clintonia, 9705–10251)

Location precision was linked to multiple participant skills or skill deficits, primarily descriptive skills and ability to locate themselves in physical space, or willingness to devote attention to these tasks. While it would have been possible for participants to record GPS coordinates for their chosen monitoring locations, none of the organizers mentioned this occurring and there are numerous reasons hikers would choose not to use GPS devices. From a scientific standpoint, the lack of precision influenced not only the ability to do expert checking on the presence of the plants reported upon at these locations, but also the scientific utility of the data.

Verifying participants' ability to identify plants was dependent on the ability of organizers to return to the locations for which data were reported. This one-time verification was important because it provided calibration data for evaluating consistency of plant identifi-

cation across participants and over time, and also helped identify appropriate locations for establishing permanent monitoring sites. The expert check for the presence of monitoring plants at observation locations could have been an easy task, were the locations adequately described. This is due in part to the influence of the alpine environment on the plants:

All of the alpine plants up there are very slow growing perennials, so it's not likely that a volunteer in 2006 would see a Labrador tea, and then I came through in the summer of 2009, and that Labrador tea had now disappeared. Or they didn't see a Labrador tea, and that Labrador tea grew up [since their visit]. (Clintonia, 14578–15021)

The issues with geographic precision were discovered during an early step in the process of verifying plant identification accuracy. The research assistant followed a painstaking process to acquire baseline data for this analysis:

I went to all of...the trail junctions from Madison Hut to Mizpah Hut, and at each of those locations, I recorded the abundance and the presence of our six Mountain Watch species. And then also, I have a running list of species that were mistaken for Mountain Watch species. I built that list off of hiking with my friends, and sometimes...I'd follow people as they did the Mountain Watch program outside of the huts, and kind of catch where they were confusing things. (Clintonia, 10365–11474)

The researchers applied the knowledge gained from this stage of evaluation to the assessment of participants' ability to identify the monitoring plants, discussed next.

Participant Skills: Botanical Knowledge

As a researcher noted, the alpine zone is a foreign environment, even for people who know plants (*Carex*). There are a number of points of confusion for non-botanists attempting to identify unfamiliar species. For example, the flowers are of tiny and difficult to see without close inspection: the blooms on a Mountain cranberry (*Vaccinium vitis-idaea*, also known as lingonberry) are only 3–8 millimeters long and often obscured by its leaves.

The organizers were well aware that plant identification was likely a problem for some

participants. As scientists with botanical training, however, the researchers needed to better understand the nature of the identification challenges for untrained individuals. A researcher observed non-experts in the field as they attempted to identify plants:

I took some of my friends out [on the trail] who weren't botanists, and kind of watched them process the identification. ... But it was definitely much different than I expected. ... I was looking at things from a very taxonomic perspective, like recognizing what family an unknown flower was in before kind of going into the guidebook, whereas my friends were looking at the color of the petals, and the shape of the leaves, and...it was probably what I should've expected, but it wasn't how I was seeing things at all. (Clintonia, 5314-6400)

Observation of non-botanists demonstrated that the challenges participants experienced with plant identification were multi-faceted.

The plants selected for monitoring were chosen to minimize confusion with identifications while also providing good phenological indicator species. Researchers found that several plants were still frequently confused with Mountain Watch monitoring species. For example, diapensia and alpine azalea were sometimes difficult to distinguish (Figure 4.9): “when they were in bloom, they were hard to confuse because the azalea was pink and the diapensia was white. But the leaves look very similar when it wasn't in bloom. ... Anything grass-like could be confused for a Bigelow's sedge” (Clintonia, 11552-12417). While individuals who participated in the alpine wildflower program learned that “sedges have edges and rushes are round, but grasses are hollow down to the ground,” most hikers never hear this mnemonic. During participant observation, it took most of the Lodge-to-Hut adventure group members several attempts at identification before they could consistently distinguish between Deer's hair sedge and Bigelow's sedge, even after carefully attending to naturalists' descriptions and explanations.

The researchers also reported that without additional reinforcement of the full set of species to monitor, observations were skewed toward more common or familiar flowers, lead-



(a) *Diapensia* with red stalks indicating the “after flowering” phenophase.



(b) Alpine azalea with pink blooms indicating the “in flowering” phenophase.

Figure 4.9: Commonly confused alpine pincushion plants.

ing to bias in the resulting data set.

We found that when it was a really open ended on what [species] you are looking for, people would gravitate towards the really charismatic species, which in the White Mountains is *diapensia*. I knew from my experience working in the huts that *diapensia* was this all-star species that was used as an example constantly by the hut croo members. So you would have these kind of skits, in which *diapensia* was a character...songs in which people, the hut croo would specifically sing about how you shouldn't walk on the *diapensia*. So people would go out with...these *diapensia*-tinted glasses, and they'd see it everywhere and kind of pass over the least well-known species, and hone in on *diapensia* or anything that looks remotely like *diapensia*. (Clintonia, 320170-33773)

While this example highlights another issue with respect to bias and data quality, it also demonstrates that hut naturalist presentations had a positive effect on participant familiarity with alpine species and served as introductory training that enhanced contributors' ability to accurately identify plants. The slow-growing nature of alpine plants, as mentioned above as an asset to data verifiability, is another reason that hut croos made a point to tell hut guests not to step on the *diapensia*.

The data collected from these systematic studies were combined with observations submitted in the prior three years. The analysis evaluated several factors, such as self-reported

certainty levels, that might contribute to accurate plant identification:

When we ran the analysis, it actually didn't matter how certain someone was. There just wasn't a statistically significant relationship between certainty of identification and a correct identification. ... We found that there is no significant relationship between flowering and identification. ... [If] something is flowering, they're more certain that they're getting it right, but they aren't necessarily getting it right. (Clintonia, 12462–17365).

This was a disheartening discovery for the organizers, who had hoped that participant's reported certainty would be a usable indicator of data quality. The fact that plants having flowers was also inadequate to predict an accurate identification was even more surprising.

These analyses provided substantial food for thought for the organizers, who began rethinking the participation protocols from a different perspective.

That's when we started having the talks about well, what are we actually asking them to do? We're asking them to do a bunch of kind of difficult tasks for someone who hasn't spent time in the alpine zone. ... We're asking them to find out where they are, to identify six flowers from what looks like hundreds of different things in front of them. ... And then we're asking them to do phenophases, which is probably a word they haven't heard before. (Clintonia, 43007–433943)

This careful data quality evaluation led to another pilot study to determine the value of redesigning the data sheets.

Science and Participation: Data Sheets Revisited and Revised

The location and identification issues that were highlighted in the prior sections led Mountain Watch organizers to revise the instruments used to record observations in the field. The primary changes to the data sheet were threefold. They started with the shift from open-ended locations to specific monitoring plots with a map and detailed descriptions for reference, plus space for self-selected locations as well. Other modifications were including lists of the known monitoring species occurring at each of the plots, and a checkbox to indicate whether the contributor had participated in a naturalist program which is intended for

future evaluation of the effects of introductory training on data quality. These changes can be seen in Figure 4.5, in which the data sheet on the right is representative of the 2009–2010 data sheets, and the left and top data sheets show the revisions made for 2011.

These revisions were tested with a pilot project in 2010 to verify that they could improve overall data quality:

The pilot study [was] looking at if the new data sheets actually did work, and we were able to show people hiking from Mizpah [Spring Hut] to Lakes [of the Clouds Hut] were providing better data than the people hiking from Lakes to Mizpah because the people from Mizpah to Lakes had the new data sheet, and were going to specific places. And the people hiking from the other direction but on the same trail, they just didn't have the focus, and their data was kind of all over the place. (Clintonia, 43944–44516)

Systematic and usability-focused evaluation of a data sheet is an established practice in volunteer monitoring projects, but is more likely to be based on heuristics than research results. This experimental design for evaluation of data quality paid off with certainty of the benefits of making further changes. The statistical analyses previously discussed also represent a more rigorous approach to appraisal of the data collection instrument than was encountered with other citizen science projects in the broader organizational field.

4.4.2 Multiple Layers of Data Collection

Another approach that organizers took to help ensure scientific data quality was the integration of the citizen science project with conventional scientific research and the use of complementary technologies. The project organizers have implemented three different methods to capture data about phenophases: citizen science volunteers (including hut guests, Adopt-A-Peak volunteers, and independent participants), trained hut naturalists, and automated data collection using plantcams. These data sources also complemented ongoing data collection by research scientists of soil and air temperature, as some of the plots at which automated data collection occurs were also designated as Mountain Watch observation sites.

Human Observation

Direct observation by humans is considered ideal, as the data require relatively little additional processing and can be directly entered into research databases by organizers or participants. Mountain Watch has leveraged two different groups to provide this type of data: volunteers and hut naturalists.

Volunteers' data is considered most subject to error due to the lack of training and participant skill evaluation, which were outside of the project organizers' capacity to implement. In addition, because the project is conducted on USFS land, regulations restricting visitor studies make it impractical to survey volunteers to learn about their prior experiences with botany and knowledge of alpine plants. Therefore, the knowledge that hikers bring to Mountain Watch participation was largely unknown and undiscoverable, which exacerbated concerns about data quality.

After years of gradual revisions to the participation protocol and data sheet, organizers felt that the 2011 version of Mountain Watch was very promising for producing good results. Nonetheless, doubt remained as to the scientific value of volunteer-based observations given the need for a fairly complex protocol:

If we can't get it [usable data] out of this year, then we will really start to question whether volunteers can get us what we need. I say it's going to work for the science; well, what's working for the science is that we have the permanent plots which the naturalists are running. (Geum 2011, 37405–37732)

Even should the volunteer-contributed data prove to be of little value, several of the organizers believed that the project was likely to continue in some fashion due to the mission-based value for education and outreach.

The hut naturalist observation, however, is likely to continue in the long term. This is in part because hut naturalists are a guaranteed resource, making observations takes relatively little time, the naturalists are trained by research staff in how to make reliable observations,

and these crew members have time specifically earmarked for these responsibilities. Hut naturalists may seem like the ideal solution to challenges with engaging hikers in Mountain Watch, but there are limits to the data that can be collected by these individuals alone. Hikers can augment hut naturalist observations with additional monitoring locations, greater aggregated frequency of observations, and opportunities for triangulation of responses.

The observation protocol for hut naturalists varied slightly from the expectations for hikers, although the data they collect were aligned with that requested of volunteers. Although the hut naturalists were expected to provide more reliable observations than casual participants, an organizer noted a caveat with respect to the skills of the crew members: “A lot of them are not trained in botanical science. Most of them are scientists...but that sometimes can be the biggest challenge, is getting them to think...of being accountable to do that [scientific observations]” (Cornus, 23741–24040). Despite these concerns, the higher likelihood of better quality data from leveraging the long-term resources represented by hut naturalists provided organizers with some certainty that with time, the Mountain Watch research goals may be met.

At the same time, an example highlighted the nature of these potential problems with confounds introduced by hut naturalists leading groups to make observations:

It just didn't make sense, because people were finding Mountain avens there, and there are no Mountain avens anywhere near there, and there's nothing that looks like a Mountain aven near there. ... [The hut naturalist] had been taking people to a site to do Mountain Watch observation...except that he had misled them about where...so they were writing down that their site was N10, but where they were was not N10, it was in a place that had Mountain avens. So they were collecting data, and they were doing a good job of it, but it showed up as being bad data in our analysis because the location was wrong, so we assumed that they had just misidentified something else as Mountain avens. (Clintonia, 49353–50817)

In this situation, the aberrant data were self-evident, the naturalist could be contacted to clarify the issue, and the data could be saved. Had this not been the case, however, a lot of

data would have been lost. Further, were the anomalies in the data less obvious, e.g., if the species reported upon were those expected for an incorrectly identified location, the issue would have gone unnoticed and introduced a less detectable form of error. To help control for this—and address several other monitoring issues—automated data collection technologies have also been introduced to complement human observation.

Complementary Technologies

As mentioned in prior discussion, the installation of plantcams served multiple purposes. In addition to providing observational data for the early and late portions of the growing season when there are few humans available to monitor plants, they also provided material for outreach through time-lapse videos, and have the potential for application to verifying volunteer-contributed data. For example, if “someone was there on the dates when the camera took a picture and is just starting to bud, and they are telling us it’s in fruit, you know that your data is no good” (Ledum, 33453–33619). Actually comparing plantcam data to hikers’ observations, however, is a more complex undertaking than it might initially appear.

The processes of implementing plantcam monitoring also required substantial effort to pilot. Organizers learned that the conditions about 4,000’ were too windy for the provided mounting hardware, and improvised with accessories made for motorcycles. They installed the cameras in numerous positions and angles to evaluate the best positioning, honed a checklist of procedures for the tricky process of downloading data, and tested the plantcams in a freezer to verify that they would operate on Mt. Washington. They also found that hikers were stealing memory cards until the devices were zip-tied shut. After working out these issues, however, the initial funding to continue this portion of the monitoring had run out. After 2010, no additional human effort could be devoted to analyzing the large volume of data produced by four images recorded daily by forty plantcams, so the number

of plantcams was reduced for 2011.

Processing image data from plantcams for this type of phenology monitoring requires human effort. This is a common focus for analysis-oriented citizen science, such as the Zooniverse family of projects, but requires customized technological infrastructure that is not readily available to most citizen science projects and too costly to customize to their purposes. The initial exploration year involved an intern to manage implementation and another intern laboriously classified the phenophases present for each species in the images, creating a large expert reference data set. Through collaboration with information science researchers, the Mountain Watch data will be tested in a game format to evaluate whether online contributors can adequately perform this step of data preparation.

Should these approaches prove fruitful, organizers identified alternate ways to record observations that would rely more heavily on technology integration: “I think as more and more people are sort of computer savvy and leaving the pen and paper behind...just taking a picture and being able to go back later and enter the data” (Ledum, 7653–7929). Combined with crowdsourced data analysis online, this approach could simplify the participation tasks to a degree that would allow considerably expansion of participation. Taking a photo of wildflowers is easy and enjoyable. Asking hikers to share their photos with researchers would also mitigate the sense that participation requires extra work on the part of contributors. After all, the hikers are on vacation.

4.5 Volunteers on Vacation

Hikers visit the White Mountains when they are on vacation. Unlike the participants in the other cases, this is a special limited-time event for them, and most people do not want to work during their vacations. The environment around them is a distraction from Mountain Watch participation tasks at all times; the challenges of the rugged terrain and breathtaking

views could make anyone forget to stop and smell the diaspensia. This section discusses the characteristics of the typical participants in Mountain Watch and the implications of this for participation and organizing. In particular, the nature of this place-based activity combined with hikers' short-term visits means that the project experiences constant churn of contributors, which organizers have attempted to combat through pervasive communication about the project. Whether or not these efforts succeed in producing scientifically useful data, AMC organizers hoped that hikers would take the message home with them in the form of individual development of personal interests, awareness, and education.

4.5.1 Hikers

Experienced hikers are typically attuned to environmental issues that impact the places they hike, in much the same way that birders care about bird conservation. If care is not taken to minimize the environmental impact of both trekking through fragile landscapes and everyday lifestyle choices, they will lose these special places. Hikers are drawn to the beauty and challenges of the White Mountains, traversing difficult trails, encountering new flora and fauna, and seeking out new experiences in the clouds. In fact, it was the expressed interest of hut guests in seeing blooming wildflowers that suggested to organizers that hikers might make good citizen scientists because every year, AMC gets calls from people asking when the wildflowers will bloom.

Most hikers encountered on the trails in the Presidentials were white, relatively young, physically fit, energetic, adventurous, and excited to be in the mountains. Observation suggested that most individuals are of middle to upper class socioeconomic status, well educated, and environmentally aware, but because AMC is not allowed to survey visitors on federally-managed lands, Mountain Watch's true participant demographics are unknown. Many hikers visiting the AMC huts are part of the larger network of AMC members, and some are members of other regional trail groups and local hiking clubs as well. This means

that hikers participating in Mountain Watch could potentially spread the word to individuals with similar interests in their personal networks, although there is little evidence as yet that this has occurred.

The more youthful demographic of hut guests than those of most volunteer groups has potential implications for the use of information technologies to support Mountain Watch. While not the primary mode of contribution at this time, online data submission may be substantially less of a challenge for Mountain Watch participants than the Great Sunflower Project's contributors. Organizers across projects had consistently reported that older individuals experienced greater difficulty with online data entry, so variation in participants' comfort using web-based data forms may be affected by generational differences in the overall participant population composition. Another organizer mentioned further incorporating mobile devices into Mountain Watch monitoring as a way to expand participation:

For this next generation that is so computer and Internet savvy, quite frankly, I think we need to be thinking now about maybe incorporating...these devices that people are going to bring into the woods. You know, it's nice that we want to think we leave all that stuff behind, but the reality is, kids these days are bringing all this stuff into the woods. (Ledum, 25143–25532)

This sentiment was echoed by another interviewee, who noted the potential for far greater accuracy in data collection by virtue of automatic recording of GPS locations and photographic evidence. This organizer also expected, however, that funding for creating this sort of a mobile application would be difficult to secure.

Since Mountain Watch participants are mostly hut guests, there are thousands of potential contributors annually. Only a small portion of these actually participate, which is similar to most citizen science projects but seems to have a stronger impact on a geographically-bounded project. In general, most hikers know less about alpine wildflowers and their phenology than participants in the other cases in this study did about the domain of the project

in which they participate. This produces a particularly challenging situation for organizers:

We have hit a wall really. ... [Hikers] are not birders, they're not so passionate...and need a place to channel their passion. These are people on vacation that we're trying to rope into doing something. ... Yet we do have a subset that is interested in conservation...and we also have a subset that are plant people. But it's a small subset. ... We thought we had a much bigger audience than we really have. We've kind of come to grips with the fact that our audience is challenging. (Geum 2011, 8047–8988)

As previously discussed, the Mountain Watch organizers have pinpointed several concerns with the observation and identification skills of hikers, but are essentially locked in to working with this population for their alpine plant monitoring. Accordingly, the organizers have focused primarily on simplifying the protocol, ongoing outreach, and communicating the alignment of project activities with hikers' personal interests: “for folks to really become engaged and stay engaged, it has to have real meaning to them in their lives and in their backyards and in their interests” (Geum 2009, 9015–9172.) Making a real-world connection between citizen science participation and personal interests is a known best practice, recommended based on the belief that more personally meaningful participation experiences produce higher quality data and greater commitment to ongoing participation (Bonney et al., 2009).

Hikers seem generally receptive to this message. Like the participants in the other cases, doing something to address larger environmental issues was perceived as meaningful to many of them. As one organizer who is in frequent contact with hikers and hut guests explained,

Sometimes these global problems seem so large that it's [citizen science] a local way of getting involved in a global problem, and you know, a lot of folks that come to these mountains do have a real connection to these mountains. They come back here regularly, whether it's yearly or bringing their kids back to the same places they went to. So to feel like they can do something to help out and to protect and...get a handle on what is actually happening up here in the mountains, it's valuable. (Ledum, 3053–4376)

This alignment of personal interests with mission has shown results for other citizen science projects discussed in this study, but it is not clear whether this is true for hikers. In this case, the hikers may care more about the landscape as a whole than the individual elements of the place, while participants in the other projects in this study frequently have a particular affinity for the specific organisms that they observe.

Although one organizer asserted that hikers are unlike birders with respect to their passion for the activities involved in monitoring projects, another organizer (who is incidentally also a birder) reported observing a direct parallel with plant-loving individuals who fit the description of listers in a general sense: “I think...people are collectors. So [they like] to collect these species, and then be able to report on them” (Ledum, 9869–10079). An organizer of a project that monitors butterflies and moths reported the same experience with citizen science participants: there are some people who simply enjoy making lists, keeping records, and tracking the species they have encountered. The primary differences across these taxa appear to be the concentration of people among the general target audience whose personal interests tend toward listing, the behavioral richness and species diversity of their preferred taxon, and the accessibility of a variety of species that can continue to fuel that interest.

Alpine plants are actually quite limited in their variety, although they may appear numerous to the untrained eye. On a brief vacation, only a few phenophases are likely to be encountered for any given species unless the individual is specifically observing this characteristic of the surrounding flora. For example, bunchberry plants observed over a single day in the altitude gain of 3,150' from Pinkham Notch to Carter Dome ranged from flowering at 2,000', to budding at 3,300', to pre-flowering with no evidence of buds at 4,800'. While most hikers are unlikely to notice these details without prompting, this further emphasizes the importance of geographical precision for Mountain Watch observations.

Despite the evidence that hikers can be trained to accurately recognize Mountain Watch

monitoring species and provide quality data on their phenophases, a substantial drawback to the project's reliance on the hut guests as contributors is the continual turnover of potential participants.

4.5.2 Constant Turnover in Participation

Given that hikers visiting the White Mountains are on vacation, their stays in the huts are relatively brief and this means that there is constant churn with respect to Mountain Watch participants. Another question that this has raised is the suitability of vacationing visitors as contributors of scientific data: "The hiking audience...they're on vacation, and they're not necessarily the right audience to try to engage in this. So we've had a lot of lessons about the audience and what level of support it takes to keep people engaged" (Geum 2009, 12351–12680). These individuals were also identified as being more likely to take a participation kit without returning data. Part of the reason that one-time visitors are less likely to follow through on contributing data was hypothesized to be related to the often difficult trail conditions:

They sort of start out with the intention of doing it, and then the trail gets harder and...the enthusiasm wanes, and they are sort of more focused on just getting done with the trail. ... It could just even be a weather day, it's a rainy wet day, it's not going to be conducive to taking out a pencil and paper. (Ledum, 7059–7929)

This claim was clearly true for the Lodge-to-Hut group, who were potentially ideal participants; for all but one particularly enthusiastic fellow, interest in continued observation tapered off over the three days as muscles grew tired and trails became tougher. Although the drop boxes for returning data sheets are available at all of the AMC facilities in the Whites, follow-through was still clearly an issue because many more kits are taken than data sheets are returned.

Organizers explained that the project would likely benefit from a higher proportion of volunteers that fit a slightly different demographic than hut guests and return to the moun-

tains more frequently. These individuals would be more likely to repeatedly provide data, and particularly at the same locations: “The Adopt-A-Peak type person, that is in a regional community that already comes to the mountains multiple times, comes to their favorite trail multiple times during the year” (Geum 2009, 30372–31069). As previously mentioned, the people who were participating outside of the hut system had provided more reliable data, sometimes with detailed reports and even photos (Carex).

One of the hopes held by project organizers for further engagement of Adopt-A-Peak volunteers has to do with the minimal level repeat participation by Mountain Watch contributors. As seen in other citizen science projects, repeated participation is consistently expected to improve data quality as well as commitment.

The majority of those participating are the hut guests. But we are hoping that the hut guests may be repeats as well...like one year we had a family that went out five times. ... So we get these pockets of enthusiasm, but...to build the recruitment, retention...which we know there’s scientific value in, is beyond our capacity at this point. (Geum 2011, 25296–26748)

Likewise, resource constraints have also stymied the adoption of the forest flower monitoring protocol and the expansion from northern forests into southern forests. These forest habitats are more accessible to many AMC members than the White Mountains, and could therefore lead to higher rates of repeated participation.

Despite these roadblocks, the development of information technologies to support broader participation has the potential to further engage the local community as well as individuals throughout the region.

I think that we could be doing a better job at reaching out to the community-level volunteers that I feel are untapped, and I think that our web interface is incrementally getting better all the time, but a real sort of community engagement is not there. So in an ideal world, I would definitely improve our website so that it is more user-friendly for engaging volunteers and having volunteers exchange information. (Geum 2009, 16867–17336)

The ability for contributions outside of the Whites via online reporting is starting to show some encouraging results on a very small scale: “We definitely have some hotspots that are non-staff, which is exciting. We have this volunteer that’s been putting stuff in off...the Kancamagus highway that we have here. We have two sites in Maine that keep coming in..utilizing the system” (Geum 2011, 26515–26908). As seen in the other cases in this study, more staff time for outreach and communication would be needed to realize the potential of this resource for engaging further participation. At the same time, the strategy of person-to-person recruitment has proven highly successful for encouraging participation from hut guests, who are surrounded by messages about the value of participating in Mountain Watch.

4.5.3 Organizing Communication Saturation

Prior sections discussed the importance of the hut-based outreach for Mountain Watch participation. As organizers noted, “When naturalists were giving a program about the alpine flowers, they tended to get more observations from that hut” (Clintonia, 17086–18559). What the interviews did not fully reveal was the saturation of this message in the AMC facilities, which became clear during participant observation.

The message that participation in Mountain Watch was valuable both for AMC’s conservation efforts and developing further scientific knowledge about the White Mountains was everywhere. Hut croos mentioned the Mountain Watch program at every evening presentation during dinner, at every hut. It has become a standard part of the croo’s daily duties, and anyone staying several evenings at the huts or visiting multiple huts would hear this message repeated at every evening meal.

Mountain Watch kits were placed in highly visible locations just inside the hut doors, typically right at the registration desk or mounted on nearby walls, where guests could hardly avoid seeing them (see Figure 4.1c). A large display at the entry of the Pinkham Notch Visitor Center took an Old West theme: “Wanted: For Flowering Along the Trail and

Turning Hikers Into Citizen-Scientists,” with accompanying mug shots of the forest flower monitoring species and an explanation of the need and goals for public participation. At Mizpah Spring Hut, flyers advertising the program were posted on the inside of the toilet stall doors in a fashion reminiscent of dormitory advertising. A large poster at Lakes of the Clouds Hut on New Hampshire’s changing seasons included Mountain Watch participation under the heading “What can you do?” The High Hut passports², new for 2011, also included a double-page spread on Mountain Watch. The AMC Field Guide to New England Alpine Summits devoted a three-page section to phenology, with two pages devoted to Mountain Watch and photos of the monitoring plants in different phenophases that demonstrate the impact of altitude on plant development. The impressively pervasive messaging at AMC facilities was another demonstration of the integration of the project into the hut culture and overall organizational embeddedness that organizers have succeeded in achieving.

Although this level of saturation of communication was clearly an important factor in participant recruitment, one organizer also noted a desire to avoid being heavy-handed with the message: “We don’t want to feel like we are pushing it on people, and it’s like they have to do this, it’s their duty as hikers to do this. But we want it to be that real curiosity, and that thing one wants to do” (Ledum, 17188–17739). This statement emphasized the importance of crafting a project that participants will want to engage in out of personal interest rather than a sense of obligation.

Other citizen science projects that are less geographically bounded and more Internet dependent have leveraged participant data and visualizations to stimulate this curiosity and satisfaction from participation. Although current digital technologies supporting Mountain Watch are still limited, making it possible for participants to find their own data on a map was intended to support participant satisfaction and lead to a stronger commitment by showing

²The High Hut passports are a similar concept to the National Park Service “Passport to Your National Parks” and National Wildlife Refuge System “Blue Goose Passport,” which encourage visits to multiple public lands through collection of souvenir stamps.

that “this is actually doing something, my little piece of this data actually did help” (Ledum, 15107–15197). The project organizers for the other cases in this study also felt that it was important to help contributors understand that however small their individual contributions, the data are valuable. By extension, the value of demonstrating how an individual’s data contribute to the larger project goals was implicitly considered important across the cases.

In addition to the value for recruitment of participants, organizers wanted their extensive outreach and education campaign to have broader impacts.

Even if you don’t submit data or anything, but you learn a couple of these species, and you say I saw this really beautiful little flower, and maybe bring a friend up here, then maybe the friend gets involved, wants to do trail maintenance or something. It’s hard to know the exact ramifications for where some of these experiences are going to lead people, but in general, it does circle through that recreation, education and conservation, and usually they get into it and they end up...finding something they’re passionate about. (Ledum, 29374–29954)

This shows the potential for engagement through personal networks and individual development in the form of new interests that are in keeping with AMC’s mission. It also reinforces the secondary goal of Mountain Watch “to educate our members about these issues and...motivate them to contribute...when they can weigh in on policy initiatives” (Geum 2009, 2401–2546).

4.6 Answering the Research Questions

Returning to the research questions, Mountain Watch had several fundamental similarities to the other cases in this study, which will be discussed further in following chapters, but also demonstrated striking differences, some of which are representative of a large segment of the wider organizational field of observation-based citizen science projects.

Virtuality was far less of a concern for Mountain Watch than the other cases in this study. This was because much of the participation takes place in a relatively bounded geographic

space and often involves group participation or instruction from a knowledgeable individual. While tools are in place to support a larger scale of participation with more remote contributors, this is not as yet a substantial portion of the observations received. Interestingly, however, the contributors who participate independently at a distance appear to provide higher quality data, which suggests that more aggressively pursuing project expansion in this area could prove valuable.

Technologies play a very different role in Mountain Watch from the other projects discussed in this study. More important than digital tools were the paper data sheets, which had more appropriate affordances for hiking conditions in the White Mountains. Instead of being a fundamental resource and a foundational requirement for launching the project, online data entry was added gradually over time with the goal of expanding participation beyond the primary contributor base. Other information technologies such as plantcams provided opportunity to collect data automatically that could be used for verification of contributed data. There are currently several hurdles that prevent that approach from being implemented, but should the resources become available, the data could be retroactively reviewed due to the multifaceted monitoring approach.

Through detailed evaluation of data quality and careful analysis, organizers identified challenges with the participation protocol that could be modified to improve data quality. In particular, the tasks of identifying locations and species were substantially simplified, and a pilot study demonstrated that this was a successful approach to supporting more reliable data collection. The evolution of the project's participation protocols can also be seen in the changes to the data sheets over time, a process that required great patience and diligence but promised to enhance data quality through greater precision and reliability.

Organizing the project has been challenging due to similar obstructions cited by the other cases, but also because of the nature of the participant population that results from

the limited geographical range of the project and its dependence on hut guests for the majority of data submissions. At the same time, organizers have bolstered these data with complementary sources, including regular observations by trained hut naturalists. Despite the trials of working within these constraints, the alignment of the scientific interests of the project with the organizational mission and potential for individual development on the part of participants has led to ongoing support in principle, if not resources. This has also resulted in a level and style of organizational embeddedness that may be unique in the larger organizational field of citizen science projects, due to the nature of the AMC facilities. These devoted resources allowed the involvement of hut crews in ongoing project outreach and saturation of project marketing.

Participation protocols have been simplified to better meet the scientific goals of the project, but unlike most citizen science projects, the reduction of choices for participants may support greater participation. As organizers became aware that the monitoring processes required participants to do too many unfamiliar tasks for their levels of skill and personal interest in the activity, they found ways to reduce task complexity without entirely removing contributors' ability to make choices about the way that they participate. These changes were reflected in modifications to the data sheet used for field data collection.

Another aspect of participation that was problematic was the fact that the hikers who contribute data are on vacation. As a result, they appear to be less committed to completing the task and turning in data, and may also find the task requirements too much like work to be an appealing leisure activity. Numerous aspects of the physical landscape can further prevent full participation, even among those who find it appealing to their personal interests, as weather and trail conditions can easily distract hikers from the participation tasks.

Scientific outcomes for Mountain Watch have been delayed, much like in the Great Sunflower Project, due to the extended revision cycles required to refine the participation pro-

to the point that hikers could contribute usable data. There were several reasons for these delays, including limited funding and personnel, plus the unavoidable influence of seasonality. A further complication inherent to phenology monitoring is that scientifically meaningful results require long-term data sets. Keeping a citizen science project operational long enough that protocols have been refined and sufficient data have been collected to produce scientific outcomes is a substantial sustainability problem for many projects. At the same time, the Mountain Watch organizers took an exceptionally rigorous approach to investigating the specific causes of problems with precision, accuracy, and reliability. This effort in itself produced scientific knowledge products that should not be discounted, as they benefit other citizen science organizers as well as the Mountain Watch project. This strategy allowed Mountain Watch project leaders to simplify the protocol and refine the participation kits to the point that data quality was demonstrably improved. Another interesting aspect of Mountain Watch was its integration with conventional scientific research at AMC. This is a fairly rare situation, but has permitted the organizers to identify several alternative modes of data collection and verification that may further support long-term scientific outcomes.

CHAPTER V

Great Sunflower Project

5.1 Know Bees or No Bees

When Gretchen LeBuhn had a little grant money left over from a research project, she considered setting up a traditional study of pollinator service. She could hire a graduate student to make regular observations of the same locations and report on how often bees visit the plants. Or she could try something different—and crowdsourcing was on LeBuhn’s mind. The increasing visibility of crowdsourcing in the media has suggested to researchers that the citizen science approach could provide large volumes of data at a relatively low cost. With her seed money, LeBuhn created the Great Sunflower Project, a citizen science project designed to collect observations in support of scientific research on pollinator service focused on plant–bee relationships. The project became an overnight phenomenon.

The project’s scientific focus was inspired by the general lack of scientific knowledge about wild bee populations. Bees are a critical link in the world’s food production system: as noted on the Great Sunflower Project website, they are responsible for every third bite of food. Although the scientific community has already documented the decline of pollinators in specific locales, little is known about pollinators across habitat types. Worldwide, gardens provide about 15–20% of the food supply, and can be especially valuable to the urban poor. Natural habitats are not common in urban areas, and may not be able to support

pollinators without human intervention to create urban gardens and restore habitats. In order to investigate these questions, it seemed only natural to enlist the public in collecting data about ecosystem health with simple observations of bees visiting garden flowers.

5.1.1 Project Description

The Great Sunflower Project (GSP) engages participants across North America to answer research questions that are important to understanding and protecting pollinator populations. It was founded in 2008 by LeBuhn, an academic researcher at San Francisco State University (SFSU). Leftover grant funding provided seed money for project startup, inspired by LeBuhn's idea to collect data at a continental scale by enlisting gardeners in a citizen science project. For her, organizing a citizen science project promised a larger and more geographically diverse data set to support her research, as well as an opportunity for public outreach and education. The project's initial research questions focused on understanding bee visitation rates across urban, suburban, and rural habitats. These research questions investigate the larger issues of pollinator service, which are a key part of ecological processes. The GSP is a young, underfunded, and technologically disadvantaged citizen science project that has shown remarkable potential and resilience despite substantial challenges.

LeBuhn carefully designed the project so that the traditional scientific observation model would be translatable to anyone's backyard, created data sheet that asked for only the minimum information that was needed, and learned how to develop a Drupal CMS website to provide educational content and forums as well as a data entry system. The GSP participation protocol asks contributors to plant specific species of blooming plants—starting with *Helianthus annuus*, the annual Lemon Queen sunflower (Figure 5.1)—and contribute observations of bee visitation according to a structured scientific protocol. Additional optional forms of participation include making phenology (life cycle) observations of the growing sunflowers while waiting for the blooms to mature enough for pollen production that attracts

bees; these data are shared with and archived by the U.S.A. National Phenology Network.



Figure 5.1: The Lemon Queen sunflower, *Helianthus annuus*.

Gardeners were the initial target community for the project because of the necessity of monitoring a particular species, which meant that participants would likely have to grow their own sunflowers. After getting the website in place, LeBuhn sent a dozen or so emails to people she thought might be able to spread the word, and then went on vacation. LeBuhn focused her initial recruiting on Master Gardeners who were likely to have a wider network of personal contacts that would include gardeners. When she returned two weeks later, 15,000 people had signed up. Although many had registered solely for the promise of free sunflower seeds and had no intention of contributing, as became evident from visitor referrals by freebie network websites and subsequent lack of participation, their level of intended engagement with the GSP was unknown at the time.

This growth pattern continued for the first two years: constantly increasing project registrations substantially outstripped LeBuhn's highest expectations despite minimal recruit-

ment efforts beyond some unsolicited attention from traditional news media. During this time, several adjustments were made to the participation protocols, data entry forms, and the removal of a major aspect of the project that initially drew so many registrants: no more free sunflower seeds. These developments will be discussed in more detail later with respect to project sustainability.

LeBuhn sought new sources of funding, but a recessionary economic environment meant that funding from grants and community foundations was scarce. Combined with the decision not to send seeds, her entrepreneurial efforts partnering with volunteers to produce and sell a calendar featuring bees resulted in enough money to support a part-time Outreach Director (Bombus). The new Outreach Director took on speaking engagements, wrote the project newsletter, prepared a participant survey, and helped with funding proposals and developing new partnerships. One of his most valuable characteristics, in addition to professional experience in both market research and adult education, is that he is a gardener. More importantly, he is a recreational gardener, not a “scientific” gardener. As a member of the gardening community, the Outreach Director has connections to and experience with institutions that engage with gardeners, and understands their interests and needs.

As of October 2010, approximately 5,000 contributors had submitted more than 40,000 observations over the project’s four years of operation. Contributions are very skewed, with one individual contributing over 300 data points in one season, but most providing only one or two samples. Observations are limited to the growing season, which varies in date range and duration across the North American continent. Supported by an open source content management system (CMS), the technology accepts data but does not offer straightforward tools for displaying it in useful ways.

Despite these challenges, the GSP sustained participation for several years before organizers were able to provide data outputs like a map of bee visitation rates. Given that the

contributor base has substantially different characteristics from and wider variability than communities of practice like birders, this success is attributed to the compelling mission of the project and charisma of both sunflowers and bees, the general attraction of which is evident in the wide array of consumer goods emblazoned with or shaped like bees and sunflowers, such as jewelry, clothing, and figurines.

The Great Sunflower Project is considered highly successful with respect to impressive volunteer recruitment in a very short period of time; over 90,000 accounts were created on the project website by its third year, although the actual rate of conversion from registrants to contributors is low. In informal interviews, organizers of other projects expressed envy at the GSP's overwhelmingly positive response from participants. The mission and aspects of the participation process and protocol clearly have broad appeal to both participants and researchers, as several other citizen science projects have subsequently adopted or adapted the observation protocol.

With limited resources for project organizing, however, the GSP has faced major challenges with funding, technology development, and providing support for participation. The budget for project operations from 2009–2011 was just over \$38,000, making the average annual operating budget around one-twentieth of the funding required to run eBird. The GSP represents a relatively new project that has survived its initial years of adjustment and is positioned for large-scale success if organizers are able to marshal the needed resources. While it could easily produce substantial growth in participation with an infusion of funding, gradual development seems more likely given the GSP's current constraints. Nonetheless, it appears that the GSP will survive the test of time well enough to produce a long-term data set that will provide a foundation for substantially broader research than was initially planned.

5.1.2 Data Collected for the Great Sunflower Project

Data collected for the Great Sunflower Project case study included the following sources:

- Interviews
- Participation and observation
- Documents

While the interviews provided the most insightful data collected for GSP, the other sources helped contextualize the participant experience and permitted data triangulation. Notably, the duration and relative scale of this project is such that there is simply less data to be collected than in the other cases: there are fewer organizers involved and the project has been in operation for a shorter period of time. One of the interesting qualities of this case for the purpose of comparison is that data collection began during the second year of the project's operation (2007), which meant that it offered a unique opportunity to observe the project's development in the early phases of organizing.

Interviews

Both of the project organizers for GSP participated in semi-structured interviews. A total of 3 semi-structured interviews lasting 60–90 minutes were held with two individuals; this is the entirety of the project's ongoing staff. Because of the ease of de-anonymization, LeBuhn's name is used throughout; the other staff person is referred to as Bombus, and both are referred to as project leaders or organizers throughout. A third individual involved in the project was a contracted database developer. I chose not to pursue an interview with this individual because involvement with the project was indirect, limited to database development, and I was already very familiar with the underlying Drupal content management system and associated database structures supporting GSP.

As part of a larger study, the interviews with project leaders were longitudinal, with two of the interviews conducted a year apart. In addition, three other interviews which were

part of a separate study touched upon partnership relationships with the GSP, providing additional context.

Additional informal interviews occurring at in-person events offered greater insight into the ongoing processes of project development. Over the course of data collection for this case, a project leader and I participated in three meetings focused on citizen science. During these events, we had opportunity to discuss the latest developments with the GSP in both direct conversations and in group settings.

Participant Observation

I participated as a contributor to the GSP for three consecutive summers, from 2009–2011. My initial reaction to learning of the project was enthusiasm—the protocol was simple enough that I knew I could easily follow it. I planted sunflowers immediately, prior to recruiting the GSP as a case study site, and therefore early field notes are sparse and data entry sessions were not recorded until the second season. I had much greater facility in initial use of the GSP website for data submission than the other cases because the site is based on the Drupal CMS . At the time, I was the administrator for five Drupal sites, which made the “logical” locations for various functionality self-evident to me, although this case study demonstrated that my ease in using the GSP website is an exception to the general experience.

Although I made a point of contributing several samples each season, following the participation protocol became admittedly dull after two years. The activities involved in contributing to the project (described later) required sitting in my front yard facing away from the street in a relatively unusual fashion, all the more noticeable because none of my neighbors have been observed sitting in their front yards at all over the last five years. My visibility during participation prompted questions from my neighbors and the postal carrier, providing an interesting experience in communicating about the project with non-participants.

I reported 10 samples across the three years of participation, with at least three observations per year, and to the best of my ability also contributed optional phenophase observations each season. While this is less participation than might have been ideal, reports from the project organizers showed that my level of participation was higher than most contributors and the constraints on my participation were typical of those encountered by others. The seasonal nature of participation in the GSP enforces a practical limit on contributing to the project, as did climatic conditions and unavoidable absences during the growing season. When contributing photos of bees visiting sunflowers became an option, I uploaded and shared 20 images with the Great Sunflower Project Flickr group, which changed my participation process for the second season to omit use of the paper data sheets to record observations (I reverted to the paper-based process in the third season).

I also periodically monitored the forums on the GSP website, which have relatively low levels of traffic. This source provided the opportunity to observe other participants' interactions with the project leaders and one another. Although very limited in scale and biased in terms of self-selection for participation, the forum postings were useful for triangulation and clearly substantiated several assertions of the organizers, particularly with respect to participant skills.

Documents

The documents used as data sources for this project included email newsletters, postal mailings, data sheets, presentation slides, fundraising products, and public-facing website content. A participant survey summary shared with contributors also provided background for understanding personal interests and experiences. The organizers reported on the results of the survey in interviews, but I did not have direct access to the detailed results. These documents showed how the GSP organizers work to promote contribution, retain participants, and diversify revenue streams.

5.2 How the Great Sunflower Project Works

The description of how the GSP works presented in this section discusses the underlying technology and contribution tasks for members of the project.

5.2.1 GSP Technologies

The Great Sunflower Project depends on information and communication technologies to organize participation. The information technology supporting the GSP, as previously mentioned, is the open source Drupal content management system. The modular system allowed organizers to pick and choose functionality to support participation. The modules implemented for GSP are primarily core features in the Drupal environment: pages, forums, and custom node types that use an add-on module for form submissions. Additional core functionality includes participant roles (e.g., administrator versus contributor), login management, password retrieval, site search, and for administrators only, image upload and tagging. The site has several pages of supporting content with information about bees and participation, with downloadable data sheets and instructions, and most of the site is publicly accessible.

An experienced programmer was hired to ensure proper database management and form checking for data submissions, which was needed to provide usable table structures for research data retrieval. The rest of the technology development and administration is handled entirely by the project organizers, notable because Drupal is a complex technology to learn and manage. Neither organizer has a background in web development. The level of technology competency required to implement and manage an open source CMS is relatively rare, as domain researchers and project organizers are infrequently trained in these skills for a variety of reasons.

In addition to the main website, the GSP utilized a Flickr photo group to enable optional

contributions of photographs documenting bees visiting flowers. Two hundred sixteen group members, representing a very small subset of contributors, have contributed 454 images, many of which do not represent images from data collection, but rather images of bees in participant gardens more generally. A Twitter account for the project was created in 2009, and although infrequently updated and limited in terms of audience, it was used to remind participants of planting timing, fundraising campaigns, and occasional reports of record bee counts (44 bees in 15 minutes as of 2011).

Starting in 2010, an email newsletter for communication with participants, called “The Buzz,” was changed from its prior plain text email format through adoption of a third-party email marketing and list manager service (Mad Mimi). This change in marketing tools improved the visual interest and appearance of professionalism for the newsletters.

5.2.2 Contribution Tasks

On the GSP website, the participation process is described as “four easy steps to participate”:

1. Sign up and plant your sunflower
2. Describe your garden
3. Watch your sunflower for 15 minutes
4. Enter your data online

The ease with which individual participants can complete these tasks is widely variable, but arguably simpler than the participation processes of many citizen science projects that are adapted from standard scientific protocols. Signing up requires creating an account on the Drupal site; it is a fairly standard account creation process, and over 100,000 individuals created accounts in the first four years of the project. Planting a sunflower requires little more than placing a few seeds into soil; however, as will be discussed later, even the most basic gardening skills cannot be assumed.

Describing a garden requires a one-time form submission for each observation location; most participants make observations in their own yards, and only at a single location. Garden description is the most complex form on the website, shown in Figure 5.2. Participants must choose from multiple choice answers for several descriptive qualities that enable a broader range of scientific analyses: garden size, amount of sun, level of development (urban, suburban, rural), fertilizer and supplemental water use, presence of bee hives, surrounding habitat, and slope of the terrain.

In addition, participants must enter location information, with required fields for city, state/province, and country. The resulting description that participants can view includes the form values and an inset Google Map with their garden location marked, which presumably allows them to verify the location resolution accuracy. Notably, location precision is substantially less important to the scientific goals of the GSP than the other cases in this study.

The core participation task is sampling bee visits, and participant instructions for 2011 were described in simple, straightforward language:

Pick a warm sunny day and if possible, sample in the morning. We recommend 9 or 10 am. Here are the five steps:

1. Set yourself up near your plant. Take a look at your garden and count how many sunflower plants are blooming. We recommend a cup of coffee or tea...
2. Focus in on one plant, count and record the number of open flowers on your plant. Don't count older flowers that might not have pollen or nectar. You can tell if a flower is old by touching the center part and seeing if your finger picks up pollen.
3. Write down your starting time (e.g. 10:00 am).
4. For each bee that visits the plant, write down its [*sic*] arrival time (e.g. 10:02 am)
5. Stop after 15 minutes have passed
6. Enter the data at: www.greatsunflower.org. After you login, look left and find "Submit Sample"

Home > Garden: AWiggins (Syracuse)

Garden: AWiggins (Syracuse) [View](#) [Edit](#)

Please tell us about your garden.

Details

Size of Garden: *

 Which best describes the size of your garden?

Amount of Sun: *

 Which best describes the sunlight in your garden?

Development: *

 Are you in an urban, suburban or rural area?

Fertilizer: *

 Do you use fertilizer in your garden?

Supplemental Water: *

 Do you water your garden?

Bee Hives: *

 Are bee hives in or near your garden? How many?

Habitat: *

 Which term describes the habitat surrounding your garden?

Slope: *

 Which best describes the terrain of your garden?

Slope Direction:

 If on or near a slope, which direction does the slope face?

Location

Name:
 e.g. a place of business, venue, meeting point

Street:

Additional:

City: *

State/Province:

Postal code:

Country: *

Current coordinates

Latitude: 43.026931

Location Map: A map showing the location of the garden in Syracuse, NY, near E Colvin St and Oakwood Cemetery.

Figure 5.2: The GSP interface for describing a garden.

The instructions also noted the option to take photographs and submit them via Flickr.

Data entry follows observation, but the initial technology with which contributors interact when making samples is a paper data sheet (see Figure 5.3), as direct data entry during participation is possible but would be cumbersome and potentially too slow for the pace of bee arrivals in an active garden. In practical terms, filling in the data sheet requires a hard writing surface, which is not a feature of most gardens, so a book or clipboard may also be needed in addition to a time-keeping device. An alternate approach for data collection is to use a digital camera to capture each bee's visit to the flower, and use the EXIF data from the images to extract the time for each visit.

Entering the bee observations sequentially online allows the participant to add up to 30 bees, one at a time, and requires filling in the type of bee and time of day, using a total of four drop-down entry fields (see Figure 5.4). Non-observation is accommodated with a checkbox labeled, "I didn't see any bees," which reduces potential data entry confusion while supporting submission of absence data. Absence data are very important for species distribution analyses, but are notoriously difficult to convince contributors to submit, according to several project organizers, because the common lay perception is that absence data are not observations and are not valuable. Finally, there is a comments field in which participants can record their notes on the weather or any other details they consider relevant.

In addition to contributing bee visitation samples, contributors can optionally report the dates of phenology events for sunflowers (date planted, first leaf, first flower, first ripe seeds, leaves dried and dead). These data are shared with another project that incorporates citizen science into the collection of long-term phenology data, the U.S.A. National Phenology Network (USA-NPN). The phenology reporting option is provided exclusively through online submission via either the GSP website or the USA-NPN website, in contrast to bee observations, which are also accepted on paper by postal mail. The USA-NPN site is substantially



The Great Sunflower Project

DATA SHEET

Your User Name: _____ Garden address (if you have more than one): _____

Please use these types of bees: Honey bee, Bumble bee, Carpenter bee, Green bee, Other or Don't know.

Month:	Date:	Date:		
Start time (e.g. 10:00 am). If you see no bees in 30 min, write no bees here.				
#sunflower plants blooming in your garden				
# open flowers on your plant with pollen				
	Time of arrival (e.g. 10:03 am)	Type of bee Not required	Time of arrival (e.g. 10:03 am)	Type of bee Not required
First bee				
Second bee				
Third bee				
Fourth bee				
Fifth bee				
Sixth bee				
Seventh bee				

Remember: Some of the most important data is when you do not see bees. Please be sure to share those data!!!

Comments (Such as what else is in bloom in your garden that's attracting bees?):


BEE GUIDE



<u>Bumble bee</u>	<u>Honey bee</u>	<u>Carpenter bee</u>
Are yellow and black (though some have red or white) and hairy.	Have gold to dark brown or black shiny abdomens with subtle stripes	Are often but not always totally black in color
Their abdomens are hairy unlike a carpenter bee.	While they have hair, the hairs are sparse.	The top of the abdomen is shiny not hairy like a bumble bee.
Are un-aggressive unless you appear to be attacking their nest	Generally smaller than either bumble bees or carpenter bees.	Have strong mandibles that they use to dig into wood where they establish nests
		Are not likely to sting unless handled

Bees: Responsible for Every Third Bite of Food
www.GreatSunflower.org

Figure 5.3: The GSP paper data sheet for 2011.



[About](#)
[Instructions](#)
[Forums](#)
[FAQ](#)

The Great Sunflower Project

Home

Search this site:

DIRECTORY

- ▷ [Join](#)
- ▷ [Learn](#)
- ▷ [Share](#)
- ▷ [Privacy](#)

AWiggins

- [My account](#)
- [Submit Phenology](#)
- [Submit Sample](#)
- [Log out](#)

Create Sample Data

Use this form to add data about bees that have visited your garden.

Date and time of observation:

Month:
Day:
Year:
Hour:
Minute:

Garden: *
Select the garden where you collected the data.

Plant observed:
Which species of flower were you observing? If you select *other*, tell us the species and variety in **comments** below.

Number of flowers:
How many flowers are blooming on the plant you observed? Count only those that have pollen - touch the center of the flower and check if your finger is dusty.

Order	Type of Bee	Time of Day
Bee 1	<input type="text" value="<none>"/> <small>Which species of bee?</small>	Hour: <input type="text"/> Minute: <input type="text"/>

[Add another bee](#)

I didn't see any bees.
Check this if you watched for the entire 15 minutes and no bees landed on the plant.

▷ [Comments](#)

Bees: Responsible for Every Third Bite of Food.

Figure 5.4: The GSP online data entry interface for bee observation samples.

more complex in terms of form submission but offers sophisticated data visualizations.

Subsequent to data entry, participants can view each individual sample record, but as of 2011 there was no personal data summary, visualizations (other than a map of sighting frequencies for 2010 across the U.S.), or other data outputs available. Although creating data outputs is a high priority for project development, it has been stymied by lack of adequate funding to develop more sophisticated reporting tools and integrate them into the CMS.

5.2.3 Meta-Contribution Tasks

Meta-contribution tasks in GSP are very limited, including assistance with creating fundraising sales items and answering user questions on open forums. A handful of volunteers have helped create the calendars that were sold to raise funds for the project. A few undergraduate student volunteers helped with data entry for observations mailed in on hard copy data sheets. Beyond this, however, the greatest constraint on engaging more volunteers in meta-contribution tasks is organizer capacity. There simply isn't enough attention to devote to volunteer management.

An additional challenge to harnessing participant enthusiasm for meta-contribution is that the tasks that need doing are not consistent, neatly structured, easily defined, independent, or uniform. In other words, the coordination and skill-specific dependencies are high for the articulation work with which GSP needs assistance. It was also noted that despite the fact that contributors are participating virtually, managing such interdependent volunteer work is difficult in a distributed environment, particularly with limited organizer experience in this domain. The challenges associated with organizing a virtual team to perform less structured tasks reinforces the intuitive notion that tasks designed to rely entirely on pooled interdependence are the most scalable choice for citizen science projects.

There is also no need for localized data review because bees are bees across the continent; bee identification has always been an optional task, in large part because the fast-moving

insects are small and can be difficult to identify, even with supporting photographs. Few lay people make a study of bee species, and unlike birds, bees do not make sounds that can be used to support visual impressions of species identifications. The data for each season have not yet achieved the scale that requires more advanced data filtering or verification approaches beyond application of heuristics to spot clearly anomalous observations.

5.3 Organizing Growth

While the institutional and organizational contexts of the other two cases in this study are described in some detail, institutional support is substantially less meaningful for the Great Sunflower Project. Although some resources were provided by San Francisco State University, they were limited enough to have little bearing on the analysis of the case. The most notable impact of the institutional environment is the observation that starting the project was a risky undertaking for an untenured academic, as LeBuhn was when the project was founded. The Great Sunflower Project offers insight into the early development stages of a new citizen science project, from developing partnerships to the challenges of unexpected scale and the associated costs of supporting technologies. These topics are the focus of this section.

5.3.1 Growing Partnerships

Over the first four years of the project, the GSP developed several partnerships, primarily related to contributor recruitment:

I'm in talks right now with the California Master Gardener program to introduce the sunflower project as a way to fulfill volunteer hours for Master Gardeners, and to actually use Master Gardeners as little agents to do the education on the sunflower project in their worlds. ... We're trying to partner with 4H clubs, they are kind of a way to reach a large audience who is interested in this topic. (Bombus, 4686–5427)

Reaching out to lifelong learning partners focused on adult education was also identified as a potentially valuable direction. Each of these partnerships focused on groups with members who are interested in gardening and agriculture, which was appropriate to the project's focus.

An additional partnership was developed as a specific support for ongoing participation with an online social network, YourGardenShow.com. YourGardenShow.com is an attractive website that describes itself as “a free online platform connecting gardeners of all experience levels with knowledge, tools and resources to inspire environmentally sustainable green spaces” (<http://www.yourgardenshow.com/about>). It specifically supports citizen science projects targeted toward gardeners, including projects focused on pollinators (GSP), allergens (USA-NPN), and phenology (USA-NPN). While the partnership does not provide additional funding, the goal is to provide a richer participation experience for project members whose primary interest is in gardening.

In-person outreach events are a key recruitment tool in the San Francisco Bay Area, where the organizers are based. A striking difference is the parallel focus on group participation opportunities in addition to independent individual participation, which seemed to be the norm. The observation protocol is suited to participation by classrooms and multiple individuals, such as family groups, who could work together to make observations. The project organizers have worked to support a more social form of engagement:

We have some curricula developed for grade school and middle school kids to observe bees, and learn how to grow and plant sunflowers. ... We're doing parent-children education about planting sunflowers and observing pollinators, and how parents can take that time to be with their kids and teach them about this. (Bom-bus, 5892–6623)

The strategy of focusing on youth and family opportunities is evident in other projects in the broader citizen science milieu, although interestingly, it is much more common in projects whose funding originates from sources emphasizing informal science education, which is not

the case for the GSP.

Loose partnerships have also formed with other citizen science projects, including pollinator-focused projects such as BeeSpotter and the Texas Bee Watchers. Several such projects are essentially spin-offs of the GSP with a localized focus or implemented as classroom projects, an unexpected broader impact of the project's appeal and relatively easy participation process. The degree to which engaging with other citizen science projects benefits the GSP is limited with respect to resources, e.g., they promote sales of the GSP's products, but may also provide an unexpected benefit of reducing some of the participant volume pressure on the GSP organizers.

Another notable partnership which has already been alluded to is the data sharing arrangement with the USA-NPN. Part of the intention of the optional phenology participation protocol was to give participants something to do while they waited for their sunflowers to bloom; this seemed a natural point of interest for gardeners. An organizer observed that the GSP had higher overall participation rates than USA-NPN relative to resources and sought out a partnership arrangement; however, participation in this part of the project has been fairly low. The low participation in phenology monitoring may be due to relatively little early communication about this part of the project and the difficulties introduced by using the unfamiliar term "phenology" with a general audience. "When we did the survey...we were interested in how strong our partnership was, and we asked...are you familiar with this phenology project? ... Less than 5% of the people who responded said yes." (LeBuhn 2010, 34100–34597) The clear lack of familiarity with the phenology monitoring portion of the project was borne out in participation; although 5% of 2500 respondents said they were aware of it, only 327 of approximately 4,000 contributors had recorded phenology data in the prior year.

5.3.2 Seed Money and Sustainability

The original participation model for the project incentivized involvement by sending out free packets of sunflower seeds to people who signed up to participate, but as later analysis of visitor referrals showed, some of these individuals were simply seeking a freebie and did not contribute data. As a result, unlike most projects, the immediate challenge was one of overwhelming (apparent) interest, rather than too few willing participants. The project had critical mass of potential participants almost from day one: “I would look at the numbers of people signing up, and go, ‘oh god, really? Another five thousand? Where did you come from?’ ... I’m both delighted and stressed by this” (LeBuhn 2009, 45487–45693).

Unfortunately, the first season was also a year of bad seeds, and the low germination rate among the free seeds (plus seed crushing by postal mail processing in Florida) meant that there was little data to be collected. The results were still promising enough to continue, especially as the list of potential contributors swelled to over 70,000 by the second year. The free seeds were not really free, of course, and the number of registrants signing up each year decimated the project budget: “This is so much more wildly successful each year than I expect that I keep blowing, completely blowing my budget, because there are so many people signing up” (LeBuhn 2009, 7257–7907). The situation led to tough decisions when the budget provided for 55,000 seed packets but 70,000 people had signed up.

In 2009, project organizers considered eliminating the free seeds, and the decision to make this change in 2010 immediately improved project sustainability and potential appeal for funders. “The biggest change is not sending free seeds. And the release of the financial burden of that is huge. It also has meant a diminished amount of data coming in” (LeBuhn 2010, 249–583). The choice not to send out seeds not only saved precious resources, but also promoted further contributor self-selection: fewer people registered without the offer of free seeds, but those who did were expected to be more likely to provide data. Combined with

entrepreneurial approaches to acquiring project funding and continued efforts to solicit funds from grantors, the change freed up enough resources to hire a part-time Outreach Director. Hiring a staff person not only improved project communications and capacity for supporting participation, but began to address the other major constraint on project scalability and sustainability: the limitations on project management posed by a single organizer running a large-scale project on the side. The organizers also quickly determined that the free seeds had contributed to the initial rush of registrants but subsequent low level of follow through; as previously mentioned, website referral information showed that the majority of the early registrants learned of the project through online networks of individuals seeking freebies.

As noted, eliminating sunflower seeds from the budget also meant that the project could present a more compelling value for investment to potential funders, further supporting project sustainability.

I wrote a lot of foundation grants, and...most of the money that I was asking for was to support sending free seeds out, and I felt like it really hindered our ability to promote the project and we didn't get a lot of funding from foundations. ... I think it's easier to write proposals that say I am doing this which is going to do all these good things for people in my program, rather than it simply means I'm going to be able to send seeds to all of my people. (LeBuhn 2010, 1533–2096)

The project organizers focused on foundation grants in part because of low success with other funding sources, such as the National Science Foundation. An issue that confronted the GSP with respect to acquiring more substantial research funding was the lack of broader acceptance of citizen science methods in the scientific research community:

I went to NSF for funding through the science part of it, not through the informal science education...and I got really positive feedback, but they won't fund it, and they want me to go to informal science education for it. ... I don't want to retrain myself as a science education person. ... I want to do ecology, and that's what I'm good at. (LeBuhn 2009, 44138–45003).

While enhancing participant experience through learning was considered a valuable output,

the GSP's primary focus is on scientific knowledge production. Unfortunately, the scientific establishment was not ready to risk funding research produced through public participation; the response to LeBuhn's grant proposal suggests that citizen science has been pigeonholed as an outreach activity rather than a method capable of producing valuable research, despite extensive evidence to the contrary (Dickinson, Zuckerberg, & Bonter, 2010). The way citizen science seems to be perceived by funders highlights a tension between available funding focused on informal science education and scientific interests which may be discouraging innovative applications of citizen science more broadly.

Instead, the project organizers had to explore other avenues for generating funding for the GSP. While challenging, the resultant diversification of project revenue sources is actually much more sustainable than the dominant grant-funded model supporting some large-scale citizen science projects. The GSP's new revenue streams took an entrepreneurial bent, in addition to solicitation of donations from participants. From 2009–2011, the organizers worked with volunteers to create calendars featuring professional quality beauty shots of bees, enhanced with additional details about each bee species and its specific pollination habits, emphasizing the connection between habitat conservation and food production. Since calendar sales flagged in 2010, another product was added in 2011: note cards featuring whimsical but biologically accurate illustrations of bees, commissioned from a professional artist, and again emphasizing the link between bees and food. The high production quality and visual appeal of these products made them an easy sell for bee and garden enthusiasts (see Figure 5.5).

Another strategy to bolster project sustainability was a return to the sunflower seeds, but this time acting as a vendor.

This year we sold seeds, and we asked people to get their own. ... Running a seed sales operation kept me busy and we are actually probably going to transition again this year. We've been approached by a seed company who is interested in



Figure 5.5: Bee-themed products sold to raise operating funds for the GSP.

handling the sales for us, and giving us a percentage of the profits from that. So I think we're going to move to that for next year, because they'll send them out in a more timely fashion than we do, and it would be nice to offload an administrative thing. (LeBuhn 2010, 918–1419)

The evolution of the seed strategy turned the necessity for individuals to obtaining seeds to participate from a project cost into a revenue stream. Committed and enthusiastic participants are likely to make their seed purchase from the partnering seed company, knowing that it will support the project's continued operations. Experienced gardeners, however, are just as likely to dry seed heads to plant the following year, a strategy for reducing personal expense that the project organizers suggested when the end of free seeds was announced.

As these new revenue sources proved funding adequate to meet immediate operating costs, over time the organizers expressed increasing confidence in the year-to-year sustainability of the project.

I feel pretty good about next year, and the project's ability to continue for however

long I feel like doing it. ... My only real frustration is trying to figure how to get enough money to do the changes that we'd like to see happen now that we have an idea of what those are. And you know, that will or will not happen. We are still getting good data, even in the state that it's in. (LeBuhn 2010, 49589–50001)

Along with finding the resources to make short-term improvements to the project infrastructure, long-term sustainability for projects like the GSP remains an open question. The primary constraint and long-term risk to project sustainability, which the organizers acknowledged as potentially problematic, has been high reliance on limited organizer capacity.

Free labor is not really free; community development and volunteer management requires organizer attention and extensive investment in communication.

My big issue is handling communications. ... We are so completely budget-limited, our volunteer management is terrible. ... I sit down and respond to emails about once a week, and I try to manage, keep up with what's happening on the forums that are set up on the web site, and...try to answer people's questions as best as I can. (LeBuhn 2009 5739–6686)

Organizers also noted that given the size of the electronic newsletter subscription list, if even a small percentage of participants reply to a newsletter by email, it would generate thousands of messages to manage. Another issue that this points to is the potential impact on participant retention when citizen science projects that are operated with minimal staffing are unable to manage communication due to unexpected increases in the scale of participation.

5.3.3 Size Matters in Organizing

The decision to launch a large-scale project was motivated by the potential to learn substantially more about pollinator service than could be accomplished through conventional science processes.

I was trying to figure out how to get a measure of pollinator service across that area, and in the traditional way, what I would do is send a graduate student up there to...try and sample how many visits that plant was getting. ... I realized that there was a potential to really increase the area that we could sample by getting

help from the community, and...it occurred to me that there was no real reason not to do this at a bigger scale. (LeBuhn 2009, 1253–2257)

This is the same motivation that many researchers who were interviewed cited as the primary reason to start citizen science projects that involve the public in data collection.

One of the remarkable things about the GSP is that it continued to gather momentum despite minimal staffing and very limited efforts for recruitment.

I actually did no outreach this year to increase participation, so everything we saw was either...some media stuff that happened but was not solicited from us, and word of mouth. So I think had I worked at trying to get more participants, I could have had double the number that we have. (LeBuhn 2009, 45105–45450)

The fast rate of growth continued even after the free seeds were discontinued; the factors supporting participant interest are discussed later. Operated by just two people, the combined time they were able to devote to the project was less than a single full-time employee equivalent. This level of staffing seems appropriate for small, localized efforts, but is exceptional here given the scale of interest and participation that the GSP has achieved and the speed at which it was accomplished.

Given the available human resources for project leadership, GSP has demonstrated the ability to coordinate the contributions of several thousand participants on par with that of larger and more established projects like eBird, which had far greater financial, human, and institutional resources at a similar stage of project development. A notable departure is that there are no full-time staff and limited organizational support. The project founder is on faculty at a university, with a 60-hour per week faculty job in addition to managing a growing citizen science project, which introduced additional constraints. Although the university setting has provided some support, it is not on par with the internal support that nonprofit organizations can devote to a new citizen science project, even with the usual limitations that these organizations typically face. The GSP could as easily have been

organized out of a home as at a university, which would not be feasible for the other cases in this study. This example demonstrates an interesting quality of citizen science projects using low-cost information technologies: when the conditions are right, large-scale participation can be generated even without the advantages of institutional resources.

5.3.4 The High Cost of Low-Cost Technologies for Participation

Distinctions between different models for free/libre and open source software development are frequently discussed as “free as in free speech” (liberty) versus “free as in free beer” (no cost). For end users like the GSP, open source software is more accurately described as “free as in free puppies”—meaning that the associated costs are high. The GSP project leaders are their own tech support; as previously noted, this is exceptional, as many domain researchers must contract out such work because the necessary technology skills are not part of their repertoires. Besides mailing seeds, “when I [have] spent money, it’s really been to manage the forms where data gets entered, because I want to make sure that there’s no way I can screw it up. Having someone who had some expertise with the best ways to data check when you have an online data entry form really was helpful” (LeBuhn 2009, 14708–14997). The choice to adopt an open source system kept operational costs low enough to fit a shoestring budget and permitted ongoing system upkeep without incurring substantial additional expense.

The main cost of using free or cheap software is not financial, but is typically reflected in usability, a known problem for the GSP. In interviews, organizers of citizen science projects with considerably larger budgets repeatedly pointed to challenges involved in combining CMS functionality with customized software that supports both scientific and participant interests, particularly data summary outputs in the form of reports and visualizations. While the Drupal CMS is easily customized to accept data, producing automated feedback for participants can be substantially more difficult. “We really need to do a website upgrade, and to get the data accessible to participants is sort of my next big goal” (LeBuhn 2010,

13889–14398). This remark was followed by the comment that sources of funding to achieve these goals had not yet been identified.

The wishlist for new functionality did not end with data accessibility. Personalizing and automating certain types of project feedback was also on the radar:

[if] you know when people plant [seeds] you could send them an e-mail saying gosh, it's been six weeks, you should be looking for a flower now, let us know when you got the flower, you know, you can automate a bunch of that stuff. ...So I think you'd be more likely to get people responding then. (LeBuhn 2010, 36015–36380)

In addition to prompting participation with personalized reminders, the ability to provide system-based feedback was also linked to potential for individual development:

One of the major challenges right now is getting data sent back...to our citizen scientists in an interactive way. Like, I'd love for them to go on the website, and be able to see what everyone else observed, and why everyone else is doing better than them, and learn how they can create a better habitat...for pollinators. (Bombus, 27516–27897)

The potential to better support participation experiences and individual development with automated prompts and system-generated feedback could substantially improve the data contribution rates and participant retention. Automated feedback, however, remained a relatively low priority compared to other more pressing concerns, such as ensuring that the data entry interface matched the current data collection protocol.

As the project's participation protocol evolved (discussed in the following section), they “redid the entry form to be more compatible with the Drupal backend, and so that works brilliantly, but the front end has maybe some usability issues. So that's another reason why I want to change it yet again, but actually make it easier, and then keep it consistent” (Bombus, 34999–35272). In addition to data entry usability concerns, discussed in more detail later, organizers also felt that the information architecture and visual design would benefit from an overhaul. Visual design is in fact a meaningful consideration, particularly

given that citizen science project organizers have expressed concern over the appearance of credibility influencing potential contributors' decision to participate. Prior research has demonstrated that attractive visual design enhances perceived credibility (Fogg et al., 2001), and a poorly designed website may suggest to potential participants that the project is not professionally managed, and therefore their contributions might be of questionable value. For projects whose contributions are motivated by personal interest in supporting the scientific goals of the project, organizers believed that credibility could be further enhanced by evidence of contributor data being put to use for the scientific goals of the project. Unfortunately, proof of the scientific utility of participation is absent on many citizen science projects' public-facing websites.

Congruent with the project leaders' main goal of scientific knowledge production, the primary investment for adapting the Drupal CMS has been ensuring that data submissions are properly recorded. As with other projects using online data submission forms, controlling input values has helped support data quality. For example, "someone entered in data that said that they saw a bee after 130 minutes, and I think what they were putting in is that it was at 1:30 in the afternoon" (LeBuhn 2010, 21445–21630). A subsequent change to the data entry form therefore requires participants to enter the time that bees were observed, rather than the elapsed time between bee sightings—clearly a point of confusion for the contributor in this example. The cost of discovering these issues was not only the developer time involved in correcting them, however, as it also revealed the loss of data that might otherwise have been usable.

Modifying the data entry form not only addressed problems such as the example above, but removed the necessity for participants to make time-based calculations, which are a potential source of error because such calculations are more challenging than most conventional arithmetic tasks; this difficulty is due to the representation of minutes in base 60 rather than

base 100 units, i.e., 60 minutes per hour instead of 100 minutes per hour. Despite the resulting nuisance factor for data entry of selecting minutes from a drop-down list of 60 items, the quality control value outweighs the minor annoyance. Several similar changes to the data entry form resulted from modifications to the observation protocol and improvements for ensuring better data quality. Organizers reported that while these changes were considered necessary from the standpoint of supporting scientific research, they caused some confusion among participants who were slow to adjust to new interfaces, particularly as the paper data sheet and the online data entry form did not entirely align at some points in the project's evolution, discussed in the next section.

5.4 Crowdsourcing Conventional Science

Like Mountain Watch and other citizen science projects, the GSP was adapted from standard scientific protocols. This choice has several implications for project development which have also been observed in a number of other citizen science projects in the broader organizational field. It involves a learning process for organizers which may require several years for a usable protocol to evolve, as well as developing an understanding of what kinds of questions volunteers can accurately answer. These interrelated challenges typically delay the production of scientific research products, which has a meaningful impact on the duration of start-up funding needed (e.g., standard 3 year grants may not be adequate in most cases) and in turn can make it difficult for a project to demonstrate sufficient effectiveness to merit funding or continued effort on the part of organizers.

5.4.1 Evolution of a Scientific Protocol for Public Participation

The way a citizen science project's participation processes and protocols evolve demonstrate strong connections between the processes of science, participation, and design. The links between these processes related to scientific interests, intended project outputs, and

perceived participant interests and skills.

The GSP organizers demonstrated great care in the design of the sampling protocols, with considerable thought given to the expected abilities of participants, which were unknown at the time. Compared to protocols for other projects with similar scientific goals, the GSP participation process is very simple. The Lemon Queen annual was initially selected as the target species—observations made on the same plant are required for comparable data—because it is native to the entire contiguous U.S., and a sunflower was chosen specifically because it would be easy to see and count bees on a large, single bloom. As a project leader explained, “I was trying to make some decisions about how I could maximize the effectiveness of the data for the science that I wanted to come out of it” (LeBuhn 2009, 33317–33487). Notably, LeBuhn had explored the participation protocols of numerous other citizen science projects in the process. She observed that many of them produce data that would be difficult to meaningfully analyze even before taking into consideration the potential confounds due to error, bias, reliability, or problems with missing data.

Where applicable, some citizen science projects attempt to reduce the usual learning curve by running pilot studies, but appears to be difficult to effectively accomplish for large-scale projects that have substantial constraints on staffing. It is essentially impossible to identify some of the problems, such as those discussed below, that can arise when a pilot project is launched on a continental scale. As a result, it took the GSP “almost 2 years of pilot to figure out the methods.” (LeBuhn 2010, 48102–48189) The first two years of the project were not intended or expected to be a pilot study, but the organizers encountered the perfect storm of overwhelming response that stretched resources to their limits combined with bad seeds, poor weather in several regions, and participant difficulties with data entry.

There were three main changes to the protocol that resulted from these early experiences: elimination of reporting temperature, reduction of the sampling time period, and accepting

data for a wider variety of plants. Temperature was removed from the data sheet, presumably due to instrument variation. The reduction of the sampling time period, from 30 minutes in the first year to 15 minutes in subsequent seasons, was largely due to the organizers' own experiences in following the monitoring protocol rather than complaints or suggestions from participants. The primary trigger for reconsidering the sampling time period was

...having a yard where I don't see bees. It drives me crazy to sit there for thirty minutes and not see anything. And so I...timed myself to see...when I started going [thinking], how much longer is this going to take? And I'm good for about ten minutes, but I can make myself stay for fifteen. I mean, I actually can make myself stay for thirty, but I'm impatient for that last fifteen minutes. (LeBuhn 2009, 35866–36431)

A related rationale for the change was to better support family participation, as the organizers had also observed that children's tolerance for the activity was about fifteen minutes. The goal for adjusting the protocol to better support family participation was generating higher levels of contribution while also meeting the educational goals of the project.

The expansion of plant species for monitoring bees is the opposite tactic from most protocol changes seen in projects based on standard scientific protocols. Usually when citizen science projects revise protocols, they reduce the options and further simplify the participation tasks to improve data quality, as was seen in Mountain Watch. While the protocol has seen modest changes over the years, the basic task is both simple and essentially the same, but the choices available to participants have expanded. Adding more species to the original Lemon Queen sunflowers provided participants additional autonomy as well as the satisfaction of being able to monitor a wider variety of plants, a very common request. The organizers worked to support the scientific goals of the project by selecting a set of suitable plants, which they have worked with seed provider partners to offer as a packaged mix. They chose plants “that there were very few varieties of, in the garden trade at least, and hope that those data would be usable. And I'm getting a lot of data on the purple coneflower.”

(LeBuhn 2009, 38904–39156)

Selecting monitoring species, an important consideration and sometimes lengthy process, impacts the contributions received and their utility for research. On the one hand, increasing the participation options means that less of the data is immediately usable for research:

Until we have enough data on those plants, I can't really use that data. I may someday be able to [use the data], and I try not to discourage people from collecting data, because I think one of the goals is to teach people just the process of science. So I...accept all data that people contribute, but I'm not going to use all of it.
(LeBuhn 2010,23184–23535)

Accepting data that are not immediately useful appears to be a valuable strategy for encouraging ongoing participation, as it treats all contributions as valuable, satisfies participants who enjoy recordkeeping, and acknowledges that data that may not be optimal for current research could have unforeseen future value. The main drawback of this approach is the required additional effort for separating out useful data from that which is not usable, but thoughtful information technology design can minimize related concerns.

The GSP organizers already see potential for wider applications of the data than were originally planned: “there’s some really valuable data that could come out of doing this for ten years. ... You could really start to get at trends and how changes in the landscape are influencing things like pollinator service” (LeBuhn 2009, 42058–42278). While the addition of plant species for monitoring yielded data that cannot be put to immediate use, it will eventually permit researchers to address new research questions comparing pollinator visitation rates for different plants. This example demonstrates that even when a project is designed to answer specific research questions, additional research opportunities may arise. It also suggests that not all changes to project protocols should be viewed as negative compromises on research goals; these changes not only satisfied popular demand, but increased the GSP’s long-term potential for contributions to scientific knowledge.

5.4.2 Asking the Right Questions

Related to the considerations that led to protocol changes described in the prior section, the GSP organizers learned another valuable lesson from these early experiences: in order to get usable data, you have to ask the right questions. Evaluating exactly which questions participants will be able to answer reliably and accurately can be difficult for organizers, and related data quality problems are likely to surface during analysis if initial data checking does not raise red flags.

The issue of which questions participants can accurately answer was identified by the GSP organizers when work on the project's first research paper began in 2010, after ironing out the initial protocol issues:

We saw that basically all of the garden sizes, until you got to very large gardens, had similar visitation rates. And that urban and rural had similar visitation rates, and suburban was lower. Which doesn't make sense to me. ... I rethought it, and realized...I had people self identify their gardens as to whether it was rural, suburban, urban, and...someone in Berkeley who would identify as they live in the suburbs, but Berkeley relative to Indianapolis is urban. So I just got a data set for the U.S. that I can use, the housing density data set...mapped all of our points onto that, and...I'm going to go back in and re-characterize the sites by rural, urban, or suburban with the external data. (LeBuhn 2010, 27991–29509)

This example of an unexpected analysis result highlighted both the types of errors that can crop up, but also the importance of considering whether participants are being asked questions for which they are able to provide valid objective answers. Protocols should therefore be examined with the specific goal of identifying questions that may elicit subjective answers to address concerns about data accuracy. Answer precision is another facet of asking the right questions that is particularly relevant to geographic data.

Geography was another aspect of describing gardens that also proved subject to errors that can be essentially eliminated through use of widely available technologies. Including latitude

and longitude is an option that GSP offers for describing locations (in addition to, rather than in place of, the required location fields), but address-based geographic resolution proved more reliable: “I had some people who put in their latitude and longitude by themselves, and they put in Soviet Union [Russian Federation]. And I’m pretty sure that since their house is in Maryland, it’s not [in Russia]” (LeBuhn 2010, 21719–23535). Even for participants who are aware that they can retrieve their location coordinates by using online services, there are numerous standards-based and formatting errors that are easily committed when individuals are unfamiliar with GPS devices and/or distinguishing between coordinate systems—skills that most citizen science project organizers would neither presume nor require.

As previously mentioned, the garden description form requires specifying city, state/province and country, something any participant should be able to accurately report. In addition, the option to omit details at the street address level helps mitigate potential privacy concerns. Since these location details were required for the analysis described above, the GSP organizers were able to integrate a housing density data set to permit the research to continue despite the subjectivity of participant classifications, as these data provide a standardized way of comparing the relative differences in human population that are used to define whether a location is rural, urban, or suburban. While these self-reported classifications into the rural/urban/suburban categories are not useful for the primary research questions, they could also be paired with the housing density data to identify any trends in the disparity between data sources, which could have a bearing on future research and project design in citizen science. This again points out that the policy of considering all data potentially valuable may have additional hidden benefits for researchers.

5.4.3 Scientific Outcomes

The prior sections identified several issues that have prevented the GSP project leaders from producing scholarly articles (at the time of writing), which included the limited

availability of researcher time and issues with data accuracy. The project has succeeded at producing a substantially larger data set than the organizers could have generated with professional researchers alone; as an organizer described it, “having data from that many sites across the US is awesome. ...We have pretty good coverage across North America, and there’s enough data to ask some really interesting questions” (LeBuhn 2010, 11799–12109). Given their commitment to producing scientific papers, it seems that achieving scientific knowledge outputs is simply a matter of time.

At the same time, certain aspects of the data contributions place constraints on the research questions to which they can be applied:

You can’t know about when people sample, relative to what’s happening in in their garden. ... We could do very different things if everybody sent in six samples a year, but we have so many...singletons and doubletons, I’ll call them, one and two samples from a garden, that just picking the max seems fair. (LeBuhn 2010, 19486–20809)

The organizers did not specify whether the situation had been foreseen, but the long-tail distribution of participation is another factor that can influence project outputs and scientific outcomes. The trend of skewed contribution rates was just as evident in the GSP data as any other online community (one contributor had provided 300 samples), and will continue to pose constraints like those mentioned above, at least until the contributor base grows to the point that adequate subsets of the data can be selected for answering questions that need six data samples per season.

Another new tactic that project organizers tried in 2011 was a concerted effort to obtain a large number of observations on two specific dates, which met with good success. The concentrated incidence of data collection allows a different type of research questions to be answered based on the temporal alignment of these data points, and also potentially increases the number of contributors providing at least two data points by emphasizing the value of

a concentrated effort and reducing contributor uncertainties about when to sample. The GSP has demonstrated that strategic communication with participants can help improve data quality and volume, which will be further discussed in the eBird case. Coordinated single-date sampling event, common in other types of citizen science such as BioBlitzes, also creates a participation experience which reinforces the sense of community that can be a motivation for some participants (Raddick et al., 2010).

5.5 Participation Experiences for Ordinary People

The GSP's first participant survey found that most participants were very happy with their participation in the project. They also found that their primary participant demographic is older adults, reinforcing the importance of good usability for online data entry. The organizers attributed the level of participant satisfaction to the match of the project mission to the participants' personal interests; they received numerous unsolicited communications from contributors describing the participation experience as empowering and uplifting. Other feedback from participants suggested that the delay involved in growing sunflowers could be negatively affecting participation, an issue that the project organizers meet head-on with regular twice-monthly newsletters during the growing season to support participant retention by providing reinforcement of the value of participation and regular reminders to participate when the flowers came into bloom. Adding the optional phenology observation protocol was also a way to help support retention while plants grew. Because the GSP organizers have yet to mount any concerted recruitment efforts, their volunteer management and communication strategies demonstrated a greater commitment to participant retention than recruitment, and strategically rewarded contributions received from the same gardens over multiple growing seasons, which may prove especially useful for research.

5.5.1 Gardeners

Birders are known for being exceptionally fanatical about making lists and sharing records of bird observations. Gardeners, however, do not have the same reputation for meticulous recordkeeping; as an organizer explained, “most gardeners aren’t really listers, and most of the people who participate have some interest in gardens, or gardening” (LeBuhn 2010, 41985–42710). Gardeners may note that certain plants have bloomed earlier or later from one year to the next, but do not necessarily keep detailed records of these changes (there are, of course, exceptions to this generalization.) Farmers, on the other hand, may have more interest in tracking these details. Current outreach related to agricultural communities has focused on youth involved in 4H, but the primary target audience has been gardeners.

The majority of survey respondents rated their satisfaction with the project highly, and the main reason for satisfaction that was cited was strikingly similar to the motivations of birders for participating in eBird: “We love bees. And we love what you guys do, and it’s really important...so they really are on board with the mission” (Bombus, 13081–13216). Despite a general lack of clarity around participants’ motivations for project participation, survey responses clearly demonstrated a strong alignment between their personal interests and the mission of the project. The organizers also recognized that some gardeners like to show off their gardens to other like-minded individuals, which spurred the development of the partnership with YourGardenShow.com. The gardener-oriented online community site features allow gardeners to monitor the development of their gardens, post images, and share their hobby with others.

These interests do not, however, easily translate to strategies that project organizers can leverage to encourage stronger commitment to contribution. Even if the GSP organizers were able to implement sophisticated technologies to support ongoing participation, there is little certainty around the types of features that might prove self-satisfying to gardeners

in a parallel fashion to the way that birders take delight in eBird features. Birders, as will be discussed later, particularly enjoy sharing data about birds and accessing that data in a variety of ways, but there is no strong evidence that gardeners feel as strongly about phenology or pollinator visitation data.

Another motivation for signing up for the project was also clear, but associated with dissatisfaction: free seeds.

When you look at participation on maps versus membership...there's these clusters of communities...where these people signed up, but yet participation is really low there. ... What the heck is going on here, that participation is so low? And then it just finally dawned on me...we're not sending free seeds anymore, so that whole motivation for being a member is not there. (Bombus, 10130–10922)

The free seeds were only one of several potential motivations to contribute, and likely one of the weakest with respect to inspiring commitment to ongoing participation. The organizers therefore felt that this was an acceptable compromise given the improvements to project sustainability and staffing. They also believed that contributors who obtained their own seeds were more likely to follow through with providing data contributions. This notion is similar to the concept of a “commitment fee,” a nominal charge sometimes required in finance (e.g., for loan processing or maintaining an unsecured line of credit) or course registration. Even very small commitment fees substantially increase follow through by weeding out individuals who half-heartedly commit to participation, and allow more effective planning based on projected participation numbers. Although incurring personal expense can prevent some individuals from participating, the project organizers said that they would send free seeds to anyone who claimed that they could not purchase their own. Since a packet of Lemon Queen sunflower seeds can be easily obtained at nationwide chain stores (including larger grocery stores) for approximately \$1.69, plus the time required to visit a store, the organizers did not feel that the expense would be a substantial imposition on participants.

Another insight from the user survey was related to the biographical characteristics of participants: “the demographic of our audience skews a little older. There are far fewer schoolchildren who participate than I thought there might be” (Bombus, 29682–29857). More specifically, the majority of participants were over 40, with a substantial proportion who were retirees. As reported by organizers of other citizen science projects more broadly, age was perceived to contribute to relatively low competency with computer use. Nonetheless, organizers were optimistic about overall trends that suggest improving fluency with technologies among older people, another reason that they pursued the partnership with YourGardenShow.com:

More and more older adults are interested in blogs, and are online a lot more than they were, say, three years ago even. So...it's [YourGardenShow.com] also good place to post pictures and share stories, and you know, they're retired and gardening, they might not have much else to do. (Bombus, 7466–7775)

Although older forum posts are no longer available on the website, a few of the earliest forum posts demonstrated that some individuals were not particularly comfortable with Internet technologies—and yet made the attempt to contribute nonetheless. Other posts included messages that stated the participant's age (60+) accompanying a complaint of difficulties using the data entry form.

While variability of participant skill is a matter of fact in citizen science, it was highlighted fairly frequently in the GSP because feedback had shown that one of the most basic skills required for participation, growing a plant that is generally considered easy to cultivate, were not always present. The lack of basic gardening skills was a substantial concern because if participants were unable to grow a sunflower, the observation tasks could not be carried out at all. The range of participant expertise on pollinators was also substantial: numerous forum posts asked for help with bee identification, while others demonstrated substantially greater knowledge of bees than the average person, using scientific names for the bee species.

As an organizer explained, “the skill base varies from Master Gardeners and beekeepers, to...amateur first-time gardeners” (Bombus, 15379–15898).

In contrast to eBird, which focused on enlisting experienced birders as contributors (only 2% of new eBird users have no birding experience whatsoever), proportionally more inexperienced would-be gardeners have signed up to participate in the GSP:

“There is a need, an identified need, for people to have some kind of coaching around growing the kinds of plants that are necessary to observe bees. ... It’s people that are really in the beginning category, that are kind of just unsure of themselves. They’ve never done it before, and all they need is just like a tiny bit of encouragement, and a little bit of luck.” (Bombus, 1188–14821)

Several forum posts confirmed the organizers’ reports that some individuals encounter problems from day one simply because they have no experience growing plants. Such hands-on skills are logistically more difficult to cultivate and support in a large-scale projects than they would be for smaller, localized citizen science initiatives where in-person trainings are a favored approach to preparing participants to contribute. Across the broader population of citizen science projects, this issue was consistently observed when interactions between organizers and participants were primarily technology-mediated.

As prior discussion revealed, the seemingly simple task of describing a garden may appear complex to some participants. Although the challenge for some participants can be partially attributed to uncertainty around some of the descriptive characteristics (e.g., direction of slope), the more common concern for garden descriptions and sample submissions was the usability of the data entry forms.

5.5.2 Demographics and Data Entry

The online data entry forms for submitting samples are relatively straightforward, and include examples and help text. This said, it is important to mention that the forms are simple for a person who is accustomed to filling out online forms. All of the issues raised

by organizers related to usability of data submission forms were based on feedback from participants by direct email, forum posts, and free text survey responses. Based on a review of the forum posts, the organizers' interpretation of participant feedback appears accurate: the users were often very specific about their complaints with respect to website usability. Some even mentioned their age as an explanation for their difficulties in using the data submission forms.

For individuals who make few online purchases and are generally unaccustomed to web-based form submission, these forms could easily seem complex and overwhelming.

One of the other things I learned from our survey is that some people, particularly older adults, have difficulty printing out the data form, and writing all this stuff in while they're observing, and taking it back, and then entering it in. ... There's a couple of places you can trip up there, so the usability of the data entry? Not great. The communication around this...could be improved. (Bombus, 23980–25524)

A more troubling issue related to website usability was echoed by both organizers, who believe that observations are being made but not reported because some participants simply find it too challenging.

I think sometimes what happens...is that folks do their observation, but they get stymied in terms of the entering process. ... Then the entry is either incomplete, or they don't do it, and although they've made the observation, we don't get the data because of that gap between the time they made the observation and actually reporting it. (Bombus, 37636–38011)

A related concern that resulted from the evolution of the protocol was indicated by participant feedback that suggested "for some people, the learning curve is longer and they kind of just got used to doing it the last way, and now there's a different way, and they think they're doing it wrong" (Bombus, 33481–34462). Several posts on the project forums by users pointed out specific inconsistencies between paper data forms and online data entry forms, which were considered problematic.

For example, the sample submission portion of the online data entry form has a small

inconsistency in naming of the bee types that differs from the language on the data entry form, which could lead to some confusion; it includes the scientific name for bumblebees (*Bombus*) but uses the common names for the other four species in the list. The online species list also offers an “unknown” option but not an “other” option despite the inclusion of “other” on the paper data sheet. While minor in isolation, issues like this can snowball into larger problems with data quality. Participant feedback about these small differences between the paper data sheet and online data entry interface suggested that they were confusing to some participants who were unsure of their performance.

In addition to the practical requirements of data collection in an outdoors setting, paper data sheets are one way that the GSP organizers try to be inclusive of “die hard retirees that want to participate, that want to feel part of the project, but yet are really still very uncomfortable with online entry, who get confused by websites” (Bombus, 38320–38520). The level of challenge such individuals encounter in submitting samples online is difficult to comprehend and fully appreciate for both the organizers and others who are comfortable with such systems; as one of them acknowledged, “I forget how unfamiliar websites are for a large part of our population, and that was...an interesting lesson to learn about communication” (LeBuhn 2010, 48102–48882). When participants’ biographical characteristics were unknown, usability was not at the forefront of the organizers’ concerns; once they became aware that the majority of contributors were older adults, these issues became a much higher priority to address as soon as funding was available.

5.5.3 Waiting for the Lemon Queen

Another part of the participation experience is growing flowers. This aspect of the participation design is in fact quite different from many other citizen science projects. Most observational citizen science projects require participants to go to a natural space and make observations “in the wild.” In the GSP, however, most contributors participate at home,

growing their own flowers in order to make observations of bees. Being able to participate from home increases the number of potential contributors, meaning that the GSP can engage a much wider audience than many other projects because participation requires no travel and little personal expense. Despite this advantage, maintaining initial participant enthusiasm continued to present challenges for organizing.

The time to flowering was identified as a specific problem, and was one of the reasons that perennials were not initially chosen (another reason was that there are essentially no perennials suited to both eastern and western regions in North America.) Growing a sunflower requires waiting, sometimes for a long time. The wait is easily long enough for a less-than-committed individual to lose interest, and many other intervening factors, both natural and personal, can influence participation drop-off in the interim. Natural interference with participation can include plant failure due to climatic patterns; predation by insects, squirrels, or deer that eliminates the plants entirely; and even weather that is intolerable or unsafe for sitting outdoors for 15 minutes due to heat or ozone levels.

For some individuals, their plants are more likely to fail due to lack of basic gardening skills, e.g., from poor choices of planting locations (despite the Lemon Queen's adaptability and hardiness), which may not seem like a concern in the eyes of contributors until they attempt to grow flowers and fail. Personal factors that reduce participation can involve any number of personal events that would prevent participation in any citizen science project, but the seasonal nature of garden-based observation also brings up challenges such as summer vacations away from home, plus the fact that most schools are not in session and therefore classroom participation is low despite the obvious match to science education and availability of curricular materials.

Even committed participants expressed impatience in their desire to contribute data. Expanding the list of flower species for bee monitoring was one way to address this complaint, as

the other plant species have different blooming dates, while the Lemon Queen sunflower can take months (up to four months in the Northeast when growing conditions are suboptimal) to mature to the point that the plant produces bee-attracting pollen and nectar. Another interim participation opportunity was added in 2009 with the optional phenology reporting component, but it too is subject to some of the same issues as sampling with mature plants.

As a result of these and other challenges, follow-through is a problem for some would-be contributors: “I get a lot of ‘maybe next year’ and it’s just sort of a funny thing. ... Once again she [a friend] said, I bought seeds you need this year, but I just never got around to getting data in” (LeBuhn 2010, 10944–11190). In a strategic attempt to address these participation issues and increase the overall volume of samples, the 2011 GSP calendar had sampling dates marked on every other Saturday in the summer months (dates are marked with sunflowers on the calendar in Figure 5.5). The project organizers also emphasized sampling on just two focal dates in project communications: “we don’t care if you don’t sample at all for the rest of the year, please do this one day. And see if that, you know, convinces people to contribute a ton of data on that one day” (LeBuhn 2010, 12801–13213). Response to the date-specific sampling campaign was excellent; according to the project newsletter, the “Great Bee Count” on July 16, 2011 produced ten times more observers entering data than any other single date since the project was founded.

Notably, the focal date approach to data collection (similar to the Great Backyard Bird Count) produced data that may have additional research applications for which the rest of the data set is less suitable. Together, this collection of issues clearly prevented some individuals who intended to participate from contributing data, but the ongoing expressions of intention to participate “next year” has promoted a volunteer management focus on retention over recruitment.

5.5.4 Organizing Participation: Retention Over Recruitment

The overwhelming success of participant recruitment with the bare minimum of effort, as previously discussed, suggests substantial potential for expansion. Recruitment therefore has not been a high priority for the GSP organizers, who were challenged to keep up with the existing contributor base, much less a larger pool of participants. Instead, the project leaders have focused their volunteer management efforts on retention of and increasing participation by existing contributors.

The project's scientific goals, and related value of data produced by repeated sampling at individual sites (which is true for all of the cases in this study), is one driver of the primary focus on retention and increasing the commitment of existing contributors. An organizer described this volunteer management task as ensuring "...that participation stays high and increases. So those who aren't...making observations, I need to make it as easy as possible for them, and bring them into the fold somehow, in terms of making observations and sending them in" (Bombus, 17511–18155). Survey feedback indicated that the most involved participants were the happiest with the project, so organizers placed further emphasis on engaging low-volume contributors more deeply. The direction of causality (and verification that it is not a spurious correlation) between participation and satisfaction is not clear, however; did the active participants contribute more because they were satisfied, or did making more contributions lead to participant satisfaction?

Responsiveness by the GSP organizers to participant feedback also led to the extension of the protocol to include additional species (Bee balm, Cosmos, Rosemary, Tickseed, Goldenrod, Purple coneflower), with minor associated changes to the data entry interface. Expanding the species selections explicitly supported retention and increased participation from already committed individuals because it produced a more satisfying participation experience, in addition to other reasons previously mentioned. This change placated some

participants, but only temporarily; even after the initial expansion of species, contributors continued to lobby for wider choice. Enthusiasm for contributing data was a clear motivator for these requests:

Some of the comments that we hear from people are that, ‘I looked at my sunflower for 15 minutes, I did not see one bee, but right next to it is the lavender plant that was crawling with bees. So why can’t I send in observations for that?’ So I want to capture and store that energy that is there, that enthusiasm for observation and collecting data. ... I see that as an opportunity to enhance, and move forward by including plants that people are already growing, that they’re excited about, that would work for the project as well. (Bombus, 18356–19054)

Despite the challenges it introduced for science processes, organizers felt that accommodating participant enthusiasm and retaining ongoing contributors was a higher priority. The project leaders also noted that the data can be sorted out on the back end, and over time the additional plant species may accumulate data sets adequate for comparison of pollinator service across plants, expanding the scientific value and applicability of the data.

Additional initiatives to support ongoing participation include collaboration with social science researchers on factors that improve retention, including commitment, thanks, gifts, and personalization. For example, in 2011, the organizers sent sunflower seeds to some participants who had contributed multiple samples in prior years. The personalized email message informing participants that seeds were in the mail also included links for recipients to indicate whether they planned to participate in 2011. Preliminary analysis of the experimental results suggested that sending seeds helped retention, even though they were sent after the optimal planting date.

Despite the variety of strategies implemented to support participant retention, the open question that the GSP organizers continued to ask is, “What can you provide to people that gives them a benefit that they want? It was sort of really an interesting thing to think about for me, and I’m not sure what the answer is to that yet. And maybe it we can’t come up

with something, I mean, it's not like you can list the bees that you've seen" (LeBuhn 2010, 41985–42710). Interestingly, project leaders for other citizen science projects focused on taxa such as butterflies, moths, and even plants confirmed that it is in fact motivating for some participants to keep a list of species they had encountered.

Why is this not the case for gardeners and bees? One reason is likely the limited number of species that most gardeners are able to identify, aside from a few of the most easily recognized species, despite the fact that many different species naturally occur across the continent. As multiple eBird interviewees mentioned, birds provide a particularly rich subject for observation due to their wide variety of behaviors, changing plumages, seasonal movements, nesting habits, and so on—these qualities do not apply to the same extent for bees, particularly as any individual's observations are usually made only in a single location. The newsletters from the GSP organizers did highlight interesting bee behaviors that participants may encounter (e.g., “pollen parties” when multiple bees simultaneously visit a bloom, or “freeloading bees” that take nectar without pollinating flowers); however, these behaviors are still fairly limited in comparison to those that birders can easily observe.

5.5.5 Encouraging Understanding and Intervention

The GSP organizers expressed an ongoing commitment to promoting individual development among contributors. Communication by the organizers is primarily through regular newsletters, which are sent twice monthly during the summer months (the same frequency at which sampling is desired.) These messages were truly informative, and included content about bees, sunflowers, and science:

When I send out newsletters, and I try to talk a little bit always about why we did some of the things...that it's important to standardize the plant, or that everybody needs to sample in the same way, or sort of what the different measures are. So just getting people a feeling for experimental design...an insight into sort of the process of science. I also try to...give people some numbers to look at, so they can

be empowered to understand what’s going on in their own yard. (LeBuhn 2009, 18042–18793)

Providing opportunities to learn about science is a common goal in citizen science projects more broadly—it is a fairly obvious avenue for individual development. In addition, professional scientists often implicitly subscribe to a deficit model of science literacy (Miller, 2001; Sturgis & Allum, 2004), and evidence has suggested citizen science participation may support “scientific thinking” in the public (Trumbull et al., 2000; Bonney et al., 2009).

Other informative content that the organizers regularly communicated were related to gardening. These messages are intended to support project participants who have little prior gardening experience while also offering suggestions for more experienced gardeners to enhance observation experiences and appeal to personal interests in gardening. The less obvious motivation for the email communications, containing suggestions for developing bee-friendly gardens, was encouraging intervention that can improve wildlife habitats and make a positive impact on environmental conditions.

I’m going to do this push for ‘adding a yard to your yard.’ That is, take a square yard of your yard or garden or lawn and change it, and convert it to be habitat. So I’ll give suggestions about what to plant, how to plant it, what it might look like, encourage people to send in photos of what they are doing. ... It would help increase the habitat for native bees and for honeybees. (Bombus, 19837–20287)

The GSP organizers saw further opportunity to enhance personal development with technology improvements to provide system-based feedback and data access. The ability to compare garden pollinator performance and identify patterns in variables across gardens, which might stimulate intervention for habitat development, was expected to fulfill a similar role that data visualizations played for eBird contributors:

These are all like the next level of little ‘aha’ moments that go on. The point of engagement would be the observation, and then the next level of that is, ‘oh, okay, so I’ve engaged, I’ve collected this data, sent it in, and here’s this feedback about

what is going on in my own garden, how I can improve habitat for pollinators?’
(Bombus, 20329–21315)

Influencing participants toward conservation or habitat improvement action is a relatively uncommon goal among strongly science-oriented projects (Wiggins & Crowston, 2012). Taking direct action is not meta-contribution, nor even a form of contribution at all, but a different type of project outcome entirely.

The organizers also mentioned that individual development was personally satisfying to both participants and organizers alike. “People gain that knowledge and have that insight about, ‘I feel bad about what’s going on on the planet, but I can actually do something in my own world to change that, and in my own thinking, and then put those principles into practice in my own garden.’ So that’s been a very powerful thing” (Bombus, 3092–3867). This quote brings up the most interesting observation that project leaders made regarding individual development: participation generated not only a feel-good response to making a contribution to science, but was also linked to a sense of empowerment. An organizer reported that numerous messages from participants conveyed the message that “I’ve been feeling so depressed about bees and these conservation issues, and I feel so good that I can do something to help. ... I’ve gotten lots of letters from people that that’s the overriding sentiment, that I felt powerless, and now I can do something” (LeBuhn 2009, 20292–20628). The sentiment expressed by these participants may be related to a sense of disempowerment over the scale of environmental damage and the limitations on what any individual can contribute in isolation. In addition, organizers hypothesized that for older adults, participation in citizen science could also provide a personally validating opportunity for meaningful contribution to the greater good that may otherwise be lacking for retirees.

5.6 Answering the Research Questions

Returning to the research questions, this case showed that many of the similarities between organizing strategies that were used by the GSP and the other case study projects were not by design, in the sense of institutional isomorphism where exemplars' practices are explicitly copied, but seem to be good responses to common challenges that may affect other citizen science projects. Surfacing and directing organizers' attention to these strategies could in fact lead to a greater level of institutional isomorphism as commonalities in organizing practices become more evident. Returning again to the research questions, the overwhelming majority of the discussion focused on the participation and organizing processes, their interrelationships, and their influences on outputs.

Virtuality received relatively little attention in interviews with project organizers, but as in other large-scale projects, is a fundamental element of the GSP. The distributed nature of participation specifically influenced the choice of measures for scientific analysis because the representativeness of samples submitted for each location is unknowable. Virtuality also influenced project design and participation due to the reliance on information technologies for data entry, although non-digital means for participating at a distance were also accommodated.

Technologies was a more prevalent theme, but information technologies were the focus, primarily in relationship to the lack of resources to make necessary improvements that can better address the needs of the participants, and the resulting challenges that contributors experienced. Data entry was more than just a hurdle for the GSP; it was clearly preventing participation or completion of participation tasks, in large part due to the biographical and demographic characteristics of participants and their difficulty using home-grown interfaces. At the same time, it is important to recognize that a small budget and lack of organizational

structure directly supporting technology development were substantial constraints on the GSP, and given these challenges, the organizers' ability to manage a large-scale project with suboptimal technology resources was no minor feat.

Organizing was a theme that cut across many of the topics in this chapter; it is clearly the primary consideration in the discussions of partnerships, institutions, and sustainability. The issue of sustainability highlighted not only the need to secure adequate financial resources to keep a project on its feet, but also the limitations of minimal staffing and over-reliance on the time of very few individuals. It demonstrated how quickly project organizers can adapt to meet funding challenges when the leaders are committed to the project and have the support of a substantial contributor base. The series of modifications to the project protocols and related topic of contributor retention was an example of a full-scale pilot, like those reported by other citizen science project organizers in interviews. The associated learning on the part of the organizers and the adaptations to the participation processes supported not only retention and contributor satisfaction, but also broadened the potential scientific outcomes and individual benefits that could be realized.

Participation processes went hand-in-hand with organizing, especially with respect to the participation protocol, and the biographical characteristics of the participant population had a meaningful impact on the technology requirements for future project development. Gardeners proved enthusiastic supporters who are dedicated to the project's mission, but many individuals were impatient with waiting for plants to grow so that they could contribute. Others lacked the basic skills to grow a flower at all, suggesting that the project has appeal to other groups besides gardeners. Nonetheless, participation seemed to produce meaningful individual development, surprisingly aligned with the idea of empowerment rather than learning or skills. Numerous concerns over usability showed how technology can dampen enthusiasm, but also the potential improvement to both participation experiences and sci-

entific outcomes that could be achieved through investment in carefully designed interfaces. Challenges that influenced the project’s scientific outcomes included learning to ask the right questions, which became evident in analysis.

Scientific outcomes for the GSP are still in progress. Data contributions had become adequate for producing scholarly articles, but issues related to protocol refinement, participation rates, and contributors’ ability to answer some of the questions posed to them delayed research outputs. As in other projects, the design of the participation protocol is central to the kind of science that can be produced, and new strategies to support year-over-year retention of participants who submit multiple samples at the same location, increased variety of plant species for monitoring, and the concurrent participation in the “Great Bee Count” represented several promising new approaches to gathering data that may expand the potential scientific contributions of the project.

CHAPTER VI

eBird

6.1 Birding for Science

The tale of the origins of eBird is the stuff of legend in the birding community. The story begins with two friends relaxing on a back porch in Ithaca, New York, discussing the future of the birds. The two men were John Fitzpatrick and Frank Gill, Executive Director of the Cornell Lab of Ornithology and Chief Scientist for the National Audubon Society, respectively. The idea they generated that evening became eBird.

Fitzpatrick and Gill had a vision to use the Internet to collect data about bird distribution and abundance on a continental (and now global) scale. The notion that average birders could contribute scientifically-useful data was a hypothesis already being tested through a series of citizen science projects at the Cornell Lab of Ornithology, with good success. With the help of a U.S. National Science Foundation (NSF) grant to make a major investment in translating citizen science to an online environment, eBird development began as the BirdSource project. The rest, as they say, is history.

6.1.1 Project Description

eBird is a popular bird monitoring citizen science project operated by the Cornell Lab of Ornithology (“the Lab”) (Sullivan et al., 2009), a nonprofit organization focused on bird conservation and research and an international leader in developing and promoting citizen

science practices. eBird allows users to keep birding observation records online:

A real-time, online checklist program, eBird has revolutionized the way that the birding community reports and accesses information about birds. Launched in 2002, eBird provides rich data sources for basic information on bird abundance and distribution at a variety of spatial and temporal scales.

eBird's goal is to maximize the utility and accessibility of the vast numbers of bird observations made each year by recreational and professional bird watchers. It is amassing one of the largest and fastest growing biodiversity data resources in existence.¹

eBird users can submit observation data by completing online checklists of birds seen and heard while birding. The system also provides tools that allows anyone to access eBird data and pursue their own questions about birds. Users can query and visualize their own data and that of others, exploring interactive maps, graphs and charts.

Contributed data are aggregated and reviewed by experts for quality when flagged as questionable by automated data filters. The data are then integrated into the Avian Knowledge Network (AKN), a public archive of observational data on bird populations across the Western hemisphere. The data archived in AKN are also deposited into larger data repositories, such as the Global Biodiversity Information Facility (GBIF). The eBird project was not designed to answer specific questions, but rather generate data that would be useful for answering a variety of research questions. eBird data have been used for policy development, conservation and land management decision-making, countless tools and reports for birders, and scientific research across several disciplines. eBird represents a mature, well supported, and technologically sophisticated project that has engaged volunteers internationally on a large scale.

eBird has become one of the best known and respected citizen science projects. Its roots go back to 1998 with the start of the BirdSource project, a collaborative effort between the Lab of Ornithology and National Audubon, after a series of previous citizen science projects

¹<http://ebird.org/content/ebird/about>

demonstrated the viability of online scientific data collection from volunteers. BirdSource began building the technological foundations to fulfill the ambitious vision of providing online tools for collecting data about birds from anyone, anywhere. An NSF grant supporting the Citizen Science Online project transformed the BirdSource project into eBird, which launched in 2002 with a focus on North American birding, as well as a number of related projects such as the award-winning All About Birds website.

Initial reception by the birding community was lukewarm, however, and contributions plateaued within a couple of years. Certain that there was far more potential than was immediately evident, Fitzpatrick instructed the project team to “turn it on its head.” And they did. When eBird version 2 was released in 2005, the birding community’s reaction was immediate and positive. Instead of asking volunteers to do “birding for science,” the project organizers created a tool that birders wanted to use for its own sake, making the scientific value of the data a secondary benefit from the contributors’ perspective. Contributions immediately took an upturn, and the graph of observation records accumulating over time became an exponential growth curve that shows no sign of slowing (see Figure 6.1.)

The substance of the changes to the system involved improved usability, marketing, and tools for birders to manage and explore bird data. The nature of the experience for birders was meaningfully changed: instead of doing science in their free time, their leisure activities were transformed into science. At the same time, two new project leaders with excellent reputations in the recreational and scientific birding communities were hired, bringing new expertise in global bird distribution and increased access to the social network of “hard-core” birders. The strategic decision to focus primarily upon serious birders as primary users of eBird made these project leaders particularly well qualified for developing solid community relationships and recruiting contributors. A third project leader with additional expertise in taxonomies was added in 2007 to help support the swiftly increasing user base and the

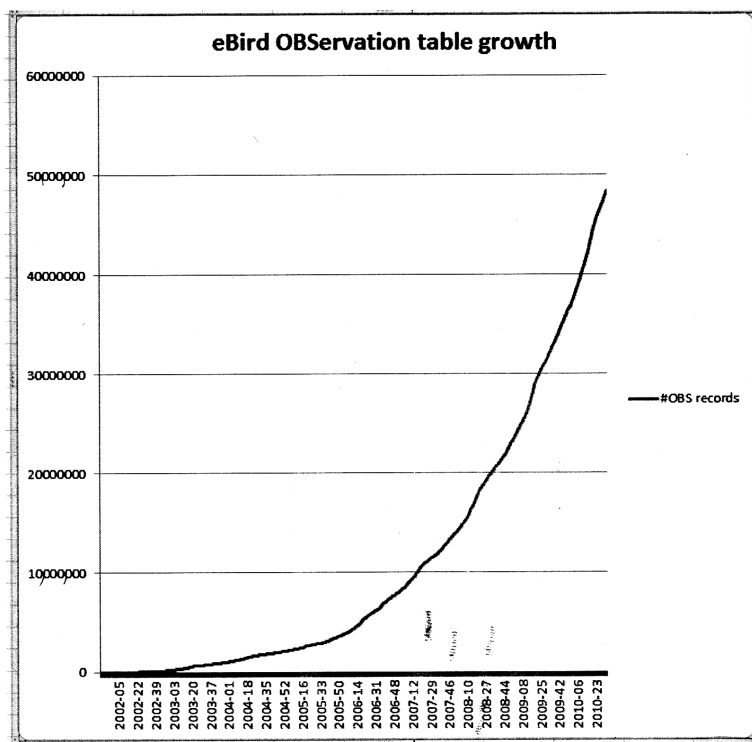


Figure 6.1: Diagram of the accumulation of records in the eBird database over time.

closely related Avian Knowledge Network project. The AKN was built on data from eBird and other citizen science projects at the Lab of Ornithology, and created a freely available collection of ornithological data sets. The dissemination of eBird data through the AKN, and the creation of the eBird Reference Data Set as a value-added scientific data product, made the data more accessible to researchers by providing a curated research-ready data set without requiring them to learn how to retrieve data from either eBird or AKN.

The second major revision, eBird Global, launched in June of 2010 and added the capacity to accept observations from anywhere in the world; previously, it could only accept observations in North America and a few other countries. Organizers reported that expert birders had expressed reluctance to adopt eBird as their primary bird record management tool because it was not able to accommodate observations from around the globe. For these individuals, whose observations are highly desirable due to their expertise and geographic diversity, the capability of storing their worldwide birding records in a single system was a

prerequisite to adoption. By the end of the year, data had already been submitted for 211 countries and over 9,000 bird species. eBird version 3 was released in November 2011, with an updated, streamlined data entry interface, improved interactive species and range maps, and the ability to embed photos with observations. These improvements supported participation through simplified data entry and exciting visualizations, while supporting data quality by allowing users to provide photographic evidence to substantiate otherwise questionable sightings.

Over time, eBird has achieved critical mass. Interviewees discussed project outputs in terms of the volume of eBird's contributions with a number of metrics: growth over time, number of observations, number of locations, number of species reported, number of countries for which data are reported, number of contributors, number of data users, and regional variations in contribution rates.

By the numbers, approximately 20,000 active data contributors at any given time are producing data that another 125,000 individuals consume (Pterodroma, 7463–7839). To date, about 45,000 participants have contributed a total of around 80 million observations, representing the largest biodiversity data set contributed by a single organization to GBIF (Dendroica, 9523–9892). The observations span 800,000 individual locations and 40,000 shared locations in 229 countries, with records for over 9,000 of the 9,969 living species listed in eBird (Fusca, 14521–14603; Dendroica, 21243–21419; Dendroica, 25431–25515). Data submissions are lowest during the month of July and peak during spring and fall migration periods, but on average, observations are added and updated at a rate of 50,000 database changes per day, and the total number of new observations being submitted ranged from 1 to 3 million per month as of 2011 (Pinicola; Meleagris, 25129–25247; Dendroica). In 2009, contributor activity added up to 80 person-years, and that was just the time spent birding (Stercorarius, 42674–42738).

eBird is widely considered one of the most successful citizen projects in existence. It has been supported by numerous grants and partnerships and recognized for its sustainability. Its adherents give glowing reviews of the software, all the while asking for more features. In exchange, users are continually contributing ever more data.

At the same time, other projects that attempt to emulate eBird have been less successful in achieving similar outcomes. This brings up the question of whether eBird's results are unique to birders and birding. The answer is, maybe.

One of the keys to eBird's success has been its strategic integration with the birding community's existing practices. Another important insight was recognizing that feedback provides crucial motivation for users by satisfying their personal interests. To the extent that projects focusing on other taxa are able to insert themselves into a community of practice and provide user-pleasing features, similar successes may be possible. The question that other projects must therefore ask is what they can provide to their contributors to motivate such high levels of engagement. From this perspective, the long history of friendly competition in the birding community (as well as its size and networks) may in fact prove unreproducible.

6.1.2 Organizational Context

As one of the largest units in the Cornell University system, with over 250 employees, the Lab of Ornithology operates at arm's length from the rest of the university, taking advantage of collaborative and infrastructural resources while maintaining an independent identity as a nonprofit organization. The Lab is comprised of a number of departments, several of which operate citizen science projects with varying levels of interdepartmental cooperation. eBird is the flagship project of the Lab's Information Science department, which focuses broadly on the use of computing technologies to support data-intensive research in ornithology through a variety of projects, including eBird.

Both organizational mission and project goals guide eBird's design and management de-

cisions. The stated mission of the Cornell Lab of Ornithology is “to interpret and conserve the earth’s biological diversity through research, education, and citizen science focused on birds.” Interviewees consistently mentioned ways that the eBird project fulfills the Lab’s mission by supporting scientific research, and also referred to the mission as a way to determine suitability, scope, or priority of projects. Conservation outcomes are also achieved through scientific research, addressing multiple aspects of the Lab’s mission. A partner organization leader noted, “each of us have a slightly different mission, and I think they complement each other in some really great ways” (Columba, 7839–7963). The goals of the eBird project are simple: to collect bird abundance and distribution data on a global scale.

6.1.3 Data Collection for eBird

The remainder of this section describes the data collection sources and strategies for this portion of the study. Data collected for the eBird case study included the following sources:

- Interviews
- Participation and observation
- Documents

These sources of data are complementary, revealing different aspects of and perspectives on the eBird project, and allowing for data triangulation.

Interviews

I conducted interviews with a total of fifteen individuals whose work relates to eBird. Eight interviews were held in person at the Lab of Ornithology, and seven were conducted by telephone or Skype. Interviews were 60–90 minutes long, based on a semi-structured interview protocol. The interview protocol provided an initial starting point for data collection, but evolved as individuals in different roles were able to address different aspects of the project. Several interviewees answered follow-up questions by email after the interview. Ver-

batim transcriptions of thirteen recorded interviews and detailed notes from two unrecorded interviews were analyzed according to procedures from grounded theory methodology.

The individuals who provided interviews were sampled on the basis of coverage and breadth, both intra-organizationally and interorganizationally. One of the project leaders served as a key informant, and recommended most of the other interviewees based on the criterion of identifying individuals with differing opinions and perspectives on eBird. Other members of the eBird team and Lab made similar recommendations that helped bring greater organizational and role diversity to the sample. Several of the interviewees are only loosely connected to the project, and therefore provided a different point of view that balanced the enthusiasm of highly involved organizers. Although attempts were made to contact individuals whose views on eBird might be less favorable, these individuals declined to provide interviews despite assurances that their perspectives would be valuable.

Pseudonym	Role	Affiliation	Birder Status
Dendroica	Director	Lab of Ornithology	eBirder
Diomedea	Portal organizer	Partner organization	eBirder
Ceryle	Manager	Lab of Ornithology	Non-birder
Circus	Organizer	Partner organization	Non-birder
Columba	Scientist	Partner organization	eBirder
Elanoides	Staff	Lab of Ornithology	Birder
Fusca	Technical staff	Lab of Ornithology	Casual birder
Meleagris	Technical staff	Lab of Ornithology	Casual birder
Otus	Portal organizer	Partner organization	eBirder
Passerina	Scientist	Lab of Ornithology	Non-birder
Pinicola	Project leader	Lab of Ornithology	eBirder
Platalea	Technical staff	Lab of Ornithology	Non-birder
Pterodroma	Project leader	Lab of Ornithology	eBirder
Setophaga	Director	Lab of Ornithology	Casual birder
Stercorarius	Project leader	Lab of Ornithology	eBirder

Table 6.1: Interviewees for eBird case study

Interviewee roles and relationships to the project and the Lab are summarized in Table 6.1; pseudonyms are based on the Latin names of bird families. All three of the project leaders, plus three technical staff and three other staff in the Information Science department provided interviews, as did two additional staff members from other departments in the

Lab. Of the eleven interviewees who are staff members of the Lab of Ornithology, three were in management positions, while others had primarily technical, marketing, or project management duties, providing a broad range of perspectives within the organization.

This sample provided perspectives from both the project leaders and technical staff for a multifaceted view of the project development and supporting technologies. A former eBird team member and a current research scientist offered complementary views, enriching the narratives of project history and current initiatives in data-intensive research. The sample included individuals at multiple levels of the Lab's organizational hierarchy, and across several departments, eliciting views that ranged from very specific and detail-oriented (e.g., from technical staff) to the proverbial bird's-eye view (e.g., from department directors). Interviewing people from partner organizations that work with the Lab in different ways and have varying relationships to the eBird project allowed me to gather external perspectives that complemented the internal point of view, including data users and portal organizers. The organizations that they represent are a mix of long-term partners and recent collaborators, which further reflected on the project's history and evolution. Since these individuals included organizers with experiences from three countries outside of the U.S., they also brought international and cross-cultural frames of reference, which were particularly helpful for contextualizing eBird's expansion to global scale data collection.

The interviews were a primary source of data for the case study. They reflect the views of eBird's organizers and leaders of partner organizations. At the same time, they are also reflective of the views of some portion of eBird participants, as several interviewees are long-term members of the birding community, use eBird regularly, and expressed very similar sentiments to those of "ordinary" eBirders observed during participant observation.

Participation and Observation

My participant observation in eBird involved birding, monitoring and participating in birding listservs, recording my own usage of eBird over time, and attending meetings at the Lab of Ornithology. This experience was an integral part of the research. While I am not an “average” eBirder, I match its new user demographics in several categories.

At the time that I began fieldwork, I had no birding experience whatsoever. Genuine participation in eBird meant that I had to learn how to bird. Learning to bird required a substantial time investment in learning how to identify wild birds, and additional investment in binoculars, field guides, and other equipment and supplies. Field notes related to these birding experiences were made periodically throughout the study.

My first eBird checklist was submitted on July 22, 2010. By the spring of 2012, I had contributed around 1,300 checklists with more than 290 species. For 2010 and 2011, I logged over 300 hours of birding on eBird checklists, equivalent to about eight 40-hour work weeks, and spent an approximately equivalent amount of time on data entry, poring over field guides, and listening to audio recordings of bird calls. While my species totals are modest for any one region, I have entered quite a few checklists, ranking #10 in the state of New York for complete checklists contributed in 2011.

Most of the active birding community communicates through email listservs. I subscribed to two local birding email lists (oneidabirds and cayugabirds) and the eBirdTechTalk Google Group. After extended lurking to learn community parlance and norms, I contributed to both oneidabirds and the eBirdTechTalk group occasionally, sending birding reports and comments on eBird features. I also beta tested and provided feedback on a new data entry interface in late 2011 and a third-party mobile application for eBird data entry in early 2012. The email listservs provided a more thorough understanding of the broader context of the birding community and contextualized the community practices that interviewees discussed.

These interactions also showed commonalities between my birding and eBirding experiences and those of others.

I maintained records of my learning process using eBird by recording data entry and site use with think-aloud sessions recorded with the Silverback usability testing software. The differences between novice and expert sessions were quite dramatic, and demonstrated substantial gains in proficiency of site use. My experiences using the eBird website provided me with direct experience in how a new user and expert user experience the site and use its tools.

In addition to interviewing several staff at the Lab, I also spent considerable time on site interacting with staff and the broader Cornell University community in the context of a graduate seminar and a reading group. I attended at least 24 meetings over the period between July 2010 to November 2011; each meeting was between one and three hours long. These visits served my intellectual interests and also brought me into close contact with the organizational culture in which eBird is situated; after months of weekly and then bi-weekly visits to Ithaca, several employees jokingly referred to me as “adjunct staff” of the Lab. As I came to appreciate the organizational complexity of the Lab as well as the broader organizational field, this close relationship with Lab staff proved a valuable resource for better understanding the eBird project.

All of these forms of participation and observation contributed to substantially strengthening the research. I experienced the common challenges and triumphs of developing bird identification skills, learned the vocabulary of birding, and came to appreciate birders’ fascination with both birds and keeping lists of them.

Documents

Internal documents Nearly 200 documents were provided directly by Lab staff, including the original grant application files and related documentation, project evaluation and user

survey reports, instructions for volunteer hotspot editors and reviewers, “canned” template messages used for communication with project contributors, and numerous other materials. These documents provided insight into the project’s development processes and the evolution of eBird.

Public documents Documents from public-facing sources such as blogs, email listserv archives, web site content, popular press and scholarly articles were also collected where pertinent to better understanding the project. These documents included blogged interviews with project leaders, articles about eBird, specific news entries and help text in the eBird website, and similar materials. These documents reflect marketing and public perspectives on eBird, user interactions with the project, and research outputs, further contextualizing internal processes and perspectives.

The following section describes the information technologies supporting participation, and the processes of contributing to eBird.

6.2 How eBird Works

The following description of how eBird functions as a technology provides a foundation for later discussion of themes such as organizing, institutional influences, and designing technology that meets the needs of both scientists and birders. This overview of the main public-facing features of eBird is followed by a brief description of the primary contribution tasks that involve members of the public as contributors to scientific research.

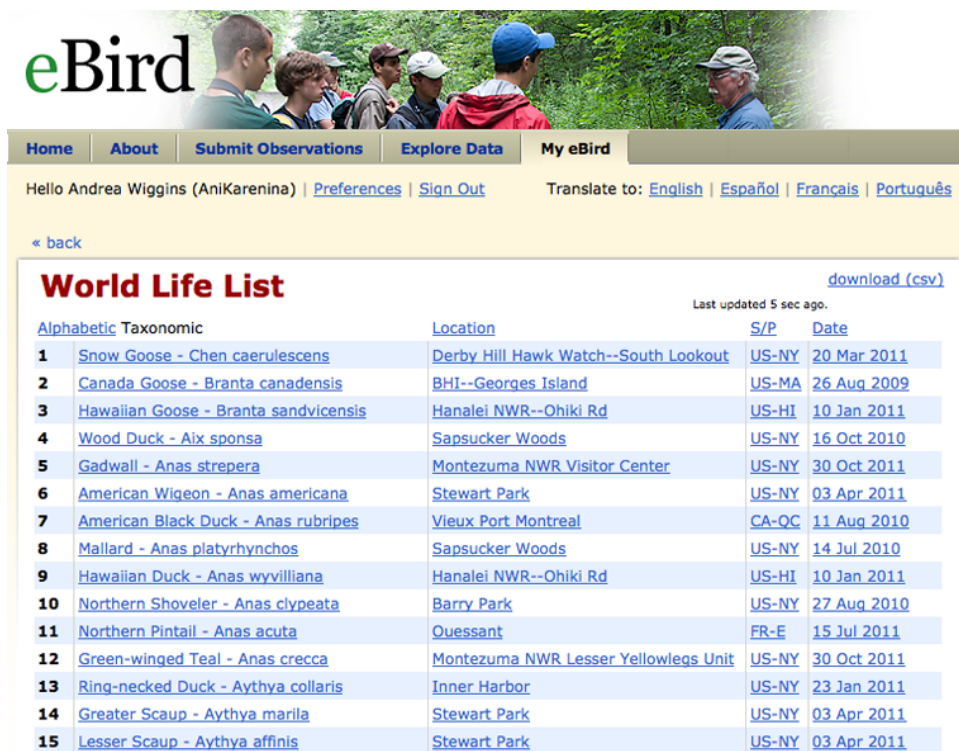
6.2.1 eBird Functionality

The primary technology supporting eBird is the custom online data submission and management software, wrapped in a content management system that provides functionality for communicating essential information for and about the project, with a presentation layer

for user interfaces. The unique core features include lists, alerts, and a wide variety of data reports, both personalized and general. It is this set of technologies that was responsible for the sharp increase in participation following the 2005 redesign of the site.

Lists

Lists are central to birders. The ways birders describe one another in the birder community usually hinges on the degree to which the birder is engaged in “listing.” Labels such as “lister” refer to the most engaged and expert birders, who are driven to collect species sightings, going to extensive lengths to see new species (however briefly) and add them to their life lists, the list of all birds they have seen in their lifetime. Maintaining a life list in a centralized online repository is one of the services that eBird provides (see Figure 6.2), and is the most basic functionality that such a data repository can provide to individuals: the ability to view their own data (Pterodroma).



eBird

Home About Submit Observations Explore Data My eBird

Hello Andrea Wiggins (AniKarenina) | Preferences | Sign Out Translate to: English | Español | Français | Português

◀ back

World Life List [download \(csv\)](#)

Last updated 5 sec ago.

Alphabetic	Taxonomic	Location	S/P	Date
1	Snow Goose - <i>Chen caerulescens</i>	Derby Hill Hawk Watch--South Lookout	US-NY	20 Mar 2011
2	Canada Goose - <i>Branta canadensis</i>	BHI--Georges Island	US-MA	26 Aug 2009
3	Hawaiian Goose - <i>Branta sandvicensis</i>	Hanalei NWR--Ohiki Rd	US-HI	10 Jan 2011
4	Wood Duck - <i>Aix sponsa</i>	Sapsucker Woods	US-NY	16 Oct 2010
5	Gadwall - <i>Anas strepera</i>	Montezuma NWR Visitor Center	US-NY	30 Oct 2011
6	American Wigeon - <i>Anas americana</i>	Stewart Park	US-NY	03 Apr 2011
7	American Black Duck - <i>Anas rubripes</i>	Vieux Port Montreal	CA-QC	11 Aug 2010
8	Mallard - <i>Anas platyrhynchos</i>	Sapsucker Woods	US-NY	14 Jul 2010
9	Hawaiian Duck - <i>Anas wyvilliana</i>	Hanalei NWR--Ohiki Rd	US-HI	10 Jan 2011
10	Northern Shoveler - <i>Anas clypeata</i>	Barry Park	US-NY	27 Aug 2010
11	Northern Pintail - <i>Anas acuta</i>	Quessant	FR-E	15 Jul 2011
12	Green-winged Teal - <i>Anas crecca</i>	Montezuma NWR Lesser Yellowlegs Unit	US-NY	30 Oct 2011
13	Ring-necked Duck - <i>Aythya collaris</i>	Inner Harbor	US-NY	23 Jan 2011
14	Greater Scaup - <i>Aythya marila</i>	Stewart Park	US-NY	03 Apr 2011
15	Lesser Scaup - <i>Aythya affinis</i>	Stewart Park	US-NY	03 Apr 2011

Figure 6.2: A life list of birds reported on eBird.

Birders' common interest in aggregating all personal bird observation records over a lifetime is one of the reasons that global expansion was important for eBird, as globe-trotting birders are less likely to be interested in managing their North American data separately from their sightings from the rest of the world. There are over 10,000 species of birds worldwide (and over 22,000 subspecies), with under 1,000 in North America and over 3,200 in South America. Data from birders who travel extensively can be very valuable, as the increasing geographic scale of the data further increases its utility.

Lists also support a number of friendly competitions to make the claim "I have more birds than you do" for some combination of time and space. eBird's list features, found in the "My eBird" and "Explore Data" sections of the website, clearly appeal to listers, both casual and the more fanatical "twitchers" for whom the expansion of the life list approaches obsession. The default lists include lifetime species numbers, and numbers of species broken down by both temporal and geographic units, e.g. the number of birds seen in Onondaga County in 2011.

Alerts

Alerts are another feature that appeal to birders. These are customizable alerts retrieved online or delivered daily by email. Alerts provide real-time information about sightings of birds species that a given eBirder has yet to record in their eBird life list or year list. The web-based summaries show the comments made on the sightings, which may include the specific location of the birds, and also whether the sighting has been verified, in the case of rarities. These alerts can be constrained to specific geographic ranges set by the user, so that someone willing to travel only within their home county can benefit just as much as a person who will hop a jet at a moment's notice to see a new bird species.

Reports

The reporting functionality that eBird provides is exceptional in the world of citizen science projects; eBird's data reporting includes both public and personal reports. The tools provide several forms of publicly available data visualization and reports, and these features are used by both eBirders and eBird's broader audience of data users. The primary categories of data tools are global range maps, bar charts for species occurrence throughout the year for any geographic region, and graphs or maps for specific species and locations, as shown in Figure 6.3.

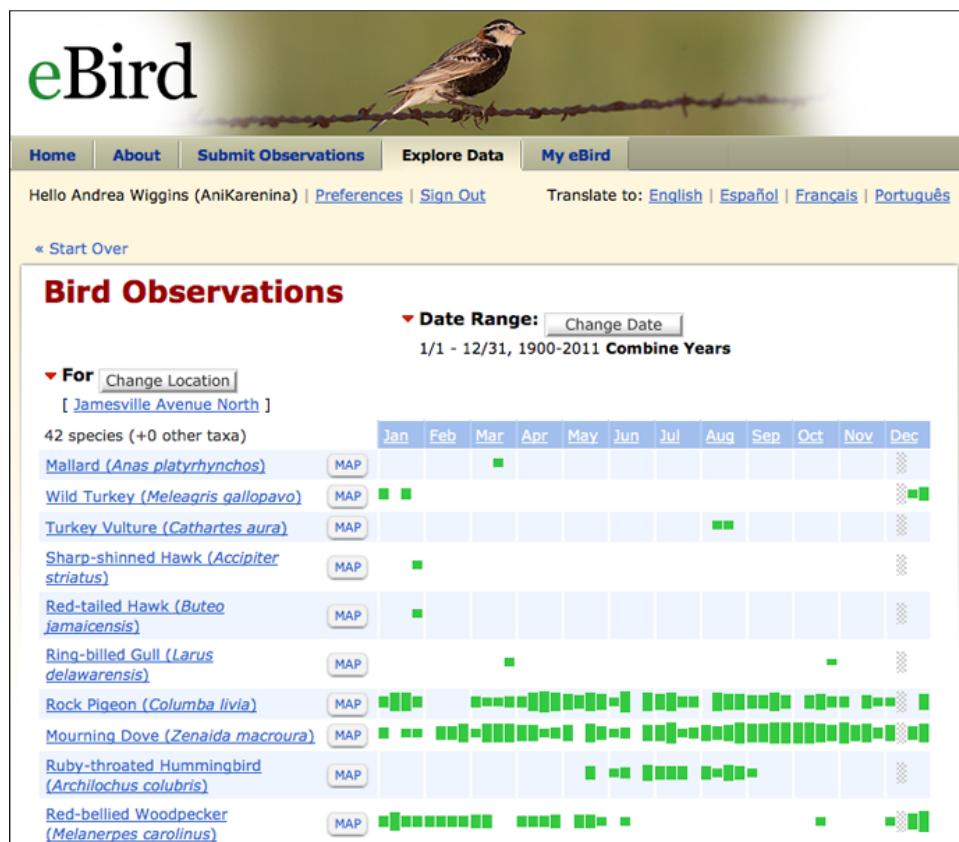


Figure 6.3: An eBird bar graph report of birds seen at a specific location.

Additional reports are inspired by existing birder practices and established procedures for sending rare sightings to other ornithological organizations. eBird also provides access to downloadable data for all reports, substantially expanding the utility of the service.

Rankings

Various contributor ranking reports also appeal to the birding community. eBird capitalizes on friendly competition by providing “Top 100 eBirders” rankings, which show personalized summaries for logged-in users. These rankings can be retrieved at multiple geographic scales from county to world, not only according to number of species reported, but also by the number of complete checklists submitted, which provides further incentive for contribution. Birders who are not able to compete on number of species can still attempt to rank highly for the number of checklists that they submit (and likely raise their ranking in number of species in the process.)

Patch and Yard Lists are another form of rankings. They are a re-imagining of the “site survey” project in which eBird had tried, unsuccessfully, to encourage participation in repeated monitoring of individual sites. The site survey served science and not birders, in that it simply requested people to make repeated observations of a favorite location, and sounded like a substantial additional commitment with respect to the regularity of observation that might be expected from a contributor. Keeping track of all the species that have appeared at a favorite birding location, called a “patch,” or in a person’s yard, however, are highly motivating to birders who enjoy seeing lists and the easy access to graphs for their favorite locations.

TrailTrackers

The eBird TrailTrackers are kiosks designed for use in wildlife refuges and nature centers. The kiosks are extensions of the eBird platform are customized to include the trails specific to the location, and allow visitors to report sightings through a simplified touch-screen interface (shown in Figure 6.5) either with or without an eBird account. The data reported anonymously through the TrailTrackers becomes part of the eBird data set under a generic

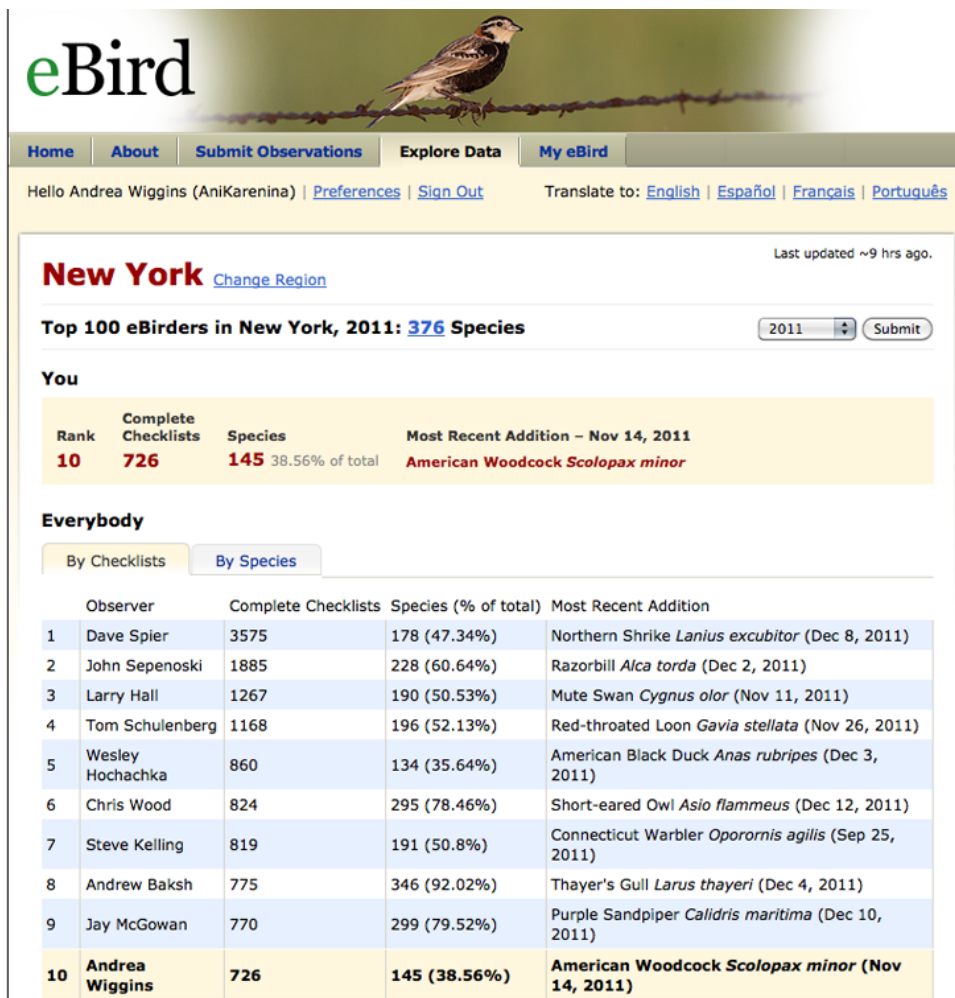


Figure 6.4: A “Top 100” report showing the rankings of eBirders in New York by number of checklists submitted in 2011.

account for the location. TrailTracker users can access recent sightings and additional information on species for that location, making it easy for a visitor to look up the most recent sightings when they begin a visit at a wildlife refuge visitor center, and to report their own sightings before they depart.



Figure 6.5: The eBird TrailTracker’s interface for assigning observations to specific locations.

The TrailTracker interface also permits users to report sightings at finer spatial granularity by allowing them to locate sightings on specific trails or specific locations along a trail. For natural resource managers, such fine-grained detail can be used for monitoring specific animals, e.g., the Great blue herons that nest in Sapsucker Woods, the wildlife sanctuary in which the Lab is situated.

Review Tools

Regional reviewers (also known as editors) have access to a special set of tools for creating filters to automatically flag unusual sightings and for reviewing flagged records. Initially, this

kind of quality review was conducted using SQL queries that checked for suspicious data. With growth of online data submission for several Lab citizen science projects, creating a review tool that worked across the projects became a priority. The data for each project has the same basic structure, making it possible to leverage the development effort required to produce review tools for any one project to support all of them. Reviewers often consult a variety of external resources (e.g., weather records) to validate observations and use tools like Excel to manage large numbers of flagged records, indicating that these tools are not comprehensive to the task at hand.

The numerous features of the eBird system support the core tasks of data contribution, described next.

6.2.2 Data Contribution Tasks

The basic process for participation via data contribution is fairly simple. A birder goes out birding, and makes a list of the birds observed, following one of several protocols (incidental, stationary, traveling, or area.) Along with the information about species encountered, details about the participation effort for the chosen protocol are also recorded: date, starting time, elapsed time, and number of observers, with some protocols requiring additional details about distance traveled or area surveyed.

The precise details of how a birder makes these records can vary substantially by individual; it is common practice in the birding community to draw up retrospective “day lists” of notable species observed during a day of birding. The traditional day list, however, does not typically contain the effort information needed for data aggregation, nor counts of birds, records of all species seen, or separate checklists for different locations. When these items are missing, it makes the data less useful for scientific research. The additional step of quantifying effort for everyday birding was described as:

...the new concept about this. The birding crowd is familiar with doing that for

Christmas [Bird] Counts, and Breeding Bird Surveys, but beyond that, pretty much the rest of the year they would just go to wherever they want, and not really think about how they're looking, what sort of effort they're putting in. (Stercorarius, 41426–42154)

By contrast, an eBirder might keep pencil and notepad at hand the entire time she is in the field, starting off by noting the time and location, jotting down every observation as they occur, and finishing up the trip by noting the ending time and distance or area covered, as appropriate. As one organizer observed, writing down observations on paper in the field is a necessary step (Circus); applications for handheld devices are not yet up to the task (although they are rapidly improving) and trusting observation to memory is unreliable. Later, the birder enters the data through the eBird web interface or uploads data using specially-formatted email messages or Excel spreadsheets. As soon as the checklist is submitted, the birder's life list and all other lists are automatically updated with new totals. Within 6 hours, the data are available through the eBird API and appear in the BirdsEye mobile app. Within 24 hours, the new observations are combined into range maps and other publicly available reports on the eBird website. The rankings for Top 100 eBirders, Yard Lists, and Patch Lists are also updated daily.

These contribution tasks are further supported by meta-contributions, defined and discussed below.

6.2.3 Meta-Contribution Tasks

“Meta” contributors support the contributions of others (Crowston & Fagnot, 2008). In eBird, these individuals are hand-selected by the project leaders as trusted contributors for a locale, and include hotspot editors and regional reviewers.

Hotspot editors review locations suggested by eBirders for potential inclusion as shared public locations where more than one person may submit data (as opposed to a private

location which is not public and not shared.) Interestingly, this terminology is slightly at odds with the usual birding community vocabulary, in which a hotspot refers to a location that is consistently particularly good for viewing birds, rather than simply publicly available. The editors are provided with a few pages of instructions, which include the naming conventions for locations in eBird. These individuals need only have local geographic knowledge to judge whether the locations submitted by other users are suitable for listing as a hotspot.

Reviewers examine data that are flagged by eBird's filters and verify the sightings, playing a crucial role in quality control. eBird's reviewer network includes approximately 400-500 individuals in North America, with a handful of international reviewers. Performing this role often requires consulting historic records for the location, plying personal local knowledge, and email exchanges with the data contributor. The task of reviewing records is sufficiently complex as to require a 29-page instruction manual and a supplemental guide to using Excel with eBird data downloads. After reviewing flagged records, the reviewers render a judgment of whether the sighting is valid or not, which is then added to the record; observations that are considered invalid still appear in the user's life list and certain data outputs where they are annotated as such.

Reviewers are local experts, selected for their expertise and willingness to volunteer their time. They have the power to not only review data but also to craft the filters that will flag unusual observations. These filters are set at the level of a geographic area (often a county, where applicable) and are at monthly intervals, because the appropriate species for a given location change throughout the year. Creating these filters can take several hours, depending on the resources that must be consulted, and require maintenance over time. Contributing as a reviewer requires a substantial additional volunteer work commitment (several hours a week for most locations), and it is clear from the delay in data review that some reviewers are periodically overwhelmed with records that require review.

The time commitment for reviewing varies, as described by a project leader:

It depends on the state. It's certainly a couple hours a week for most places. That's not counting any filter creation or anything. To do a good filter can take a couple hours or many hours, depending on...how much you have to look up, and how much can you do off the top of your head. (Stercorarius, 25995–26310)

The detail-oriented reviewing and editing tasks represent “a lot of fiddly work going through those flagged records and dealing with hotspots” (Otus, 12267–12686). At the same time, eBird project leaders hope that reviewers will go beyond the minimal commitment of reviewing flagged records: “we like to encourage people to use some other tools to help us look for problems, like misidentification of goshawk is a problem nationwide” (Stercorarius, 26422–26752). While no one discussed the motivations of these meta-contributors, besides the satisfaction of being recognized as a local expert, it stands to reason that their personal interests are likely to be somewhat different from those of average eBirders.

6.3 Organization and Organizing

eBird's organizational context and project history laid the foundation for its current leadership structure, the staff who are responsible for organizing participation. These factors also led to a strong sustainability plan to ensure eBird's long-term operation. The topics of project leadership and sustainability are the focus of this section.

6.3.1 Project Leadership

The BirdSource project was initially staffed with four employees who were internal hires from departments within Cornell University. The staffing strategy changed as the project developed, grew, and became eBird. It now functions as a distributed team, with the employees working on eBird spread out across the U.S. Employing a virtual project team not only permitted the Lab to retain premiere birders as project leaders, but also helped the

project leaders maintain stronger regional links to the local birding community and facilitated broader outreach in North America.

Strategically, this has proven valuable. By hiring well-known and respected birders to manage the project, who have a strong sense of the birding community's needs and desires, the project also benefits from the associated positive reputation of these project leaders. Part of creating a mutually-agreeable arrangement for the project leaders was ensuring that they are able to work virtually, and could take time to continue their professional birding activities. These additional activities, such as leading birding tours and editing popular birding magazines, continued to reinforce the positive reputation of eBird through the visibility of its project leaders.

Internally, the eBird project leaders work with database administrators, an interface designer, and a web applications developer to implement site development decisions. The eBird team also includes a research statistician who works with data modeling based on eBird data, and the Avian Knowledge Network staff, most of whom are assigned partial time to AKN as well as other projects. Only the web applications developer is employed to work on eBird alone; the others split their time between eBird and other projects. As a result, the project's official staffing includes only about 4–4.5 full-time equivalent employees, depending on the funding available and development projects in progress. Within its organizational context, the eBird team is also unique in being relatively self-contained, meaning that all the dedicated full-time and part-time staff who are employed to work on eBird are part of the Information Science department at the Lab. Other citizen science project teams at the Lab include members of multiple departments. The project receives additional support from other Lab departments as needed, such as promotion from the Marketing department, press releases from the Communications department, and evaluation from the Program Development and Evaluation department.

With respect to organizational human resources, a broader skill and knowledge base has been developed with lower direct investment by the project, partly because most eBird staff are assigned to the project part-time. eBird has also received more resources than other citizen science projects at the Lab, causing some envy from other groups and leading to the perception that “eBird is the favored child” (Ceryle, 11143–11338). The intellectual resources available to the project are impressive; eBird draws upon the expertise of biologists, statisticians, and computer scientists, plus the complementary domain knowledge areas of each of the three project leaders. This is achieved primarily through interorganizational partnerships.

While resources are always in limited supply, organizers also remarked upon how effectively eBird has operated with a relatively small staff. This observation extends from the eBird project team to portal organizers, who are paid by other organizations to promote and use eBird. One individual observed that the entire North American continent is managed by a total of about six staff positions, spread across multiple organizations: “that’s what’s really exciting to me, is that you can do this kind of thing with such a small team” (Otus, 33414–33532).

In addition to the project’s official staffing, of course, the larger network of contributors is a substantive part of the eBird project organization, and eBird portal organizers play a special role. They act as local ambassadors for eBird, working with the project leaders to create a localized or organization-specific eBird portal. The customizability provided the portal organizers with simple branding and access to a content management system for dissemination of customized materials and content of local or organizational interest, while hardware is managed by the eBird team and all data are stored the central eBird database.

Portal organizers work on customization, which includes tasks like language translation and the creation of instructional materials suitable to their country, local region, or mem-

bership. They also recruit contributions to their portal, plying their own personal social networks to enlist birders as eBird users. Most eBird portal organizers are employees of ornithological organizations, and organizing eBird contributions through their local portals is a part of their job. Although portal organizers do not have decision-making power with respect to the eBird system, they control several aspects of their portals. The portal organizers' work supports eBird's expansion and they are evangelists who play an important role in extending the project's reach: "another thing I do is promoting eBird, and making sure that the [portal] website is up to date, and answering questions that participants have, encouraging participation, helping people upload data" (Otus, 2355–2955). In some cases, portal organizers work in close collaboration with the eBird project leaders; for example, an individual from a partner organization is responsible for the French language translations. These portal organizers effectively extend the human resources available to eBird by adding to the leadership and coordination capacity at no additional cost to the project.

eBird's success has also yielded substantial funding to support the project, with over \$6 million in grant funding since the project's inception (Dendroica). Although this is a large figure, the grant awards were received and expended over a period of ten years, and expectations are high for a well funded project. Initially, the Lab provided internal venture capital funding to support the original BirdSource project development, which was a particularly valuable benefit of starting the project in a unique institutional environment. Since repaying the internal debt with grant funding from the NSF, the Information Science department developed a project sustainability plan which has been recognized for its excellence (Maron, Smith, & Loy, 2009).

To support eBird's long-term viability, a variety of revenue sources support ongoing project maintenance. The assurance of project sustainability is important for convincing top-level contributors and data consumers to rely on eBird as an authoritative data source

and as a data management tool, and even more so as an increasingly longitudinal data set is developed. Developing a long-term data set brings up questions of project sustainability, which are addressed next.

6.3.2 Project Sustainability

eBird’s model for fiscal sustainability has been internationally recognized for its excellence by external groups, who highlighted this structure as a model for sustaining digital resources (Maron et al., 2009). The general approach can be summarized very simply: “We use NSF money for innovation, and use other resources for sustainability” (Dendroica, 7718–7971). A similar funding model is in place for the AKN, although staff noted that “the project funding has declined recently, so we’re sort of in a maintenance mode now” (Meleagris, 6764–6866). In spite of current resource constraints, this model for long-term sustainability kept the AKN active and available to researchers.

eBird’s revenue sources included sponsorships, portal software licensing, endowment payouts, and kiosk fees for the eBird TrailTrackers. These income streams covered the costs of 4.25 full-time employee equivalents, including the project leaders, a web developer, department administrators, and a database administrator (all partial time with the exception of the web developer.) The project’s non-personnel expenses included hosting and technology costs, as well as overhead paid to the Lab of Ornithology to support organizational infrastructure (Maron et al., 2009). This funding structure meant that eBird did not rely directly on grant funding (which was used to start the project), but other projects in the Information Science department support employees who spend partial time on eBird, which provided an indirect subsidy that helped minimize project staffing costs.

The revenues generated by software licensing and kiosks represented a form of franchising that promotes eBird to new audiences, serving multiple purposes. Approximately 30 customized eBird portals brought in an initial set-up fee and an annual maintenance fee; a

similar model with an initial set-up fee and annual service fees applied to approximately 40 kiosks that are rented to nature centers. Sponsorship funding supplemented these revenue streams, although it was considered vulnerable to economic pressures. Notably, the organizational infrastructure offered by the Lab of Ornithology has allowed the project to operate at a deficit at times due to start-up costs for new developments, such as the kiosks, but these loans must be repaid to the institution to maintain the good standing of the Information Science department for similar future investments by the organization. This internal venture capital arrangement was a substantial advantage to the eBird project as it expanded over the years.

A report by the ITHAKA research group on sustainability of digital resources highlighted several broader implications from eBird's example (Maron et al., 2009). They noted that successful engagement with contributors is a result of deep understanding of user needs and interests, but rapid shifts in strategy may be necessary to maximize value, a lesson learned from eBird's resounding success in its version 2.0 release as compared to its initial reception. The sale of customized services helps support open access to the data resources, which is a central mission of the project, and notable because in general, successful funding models for open access repositories are notoriously challenging to arrange. The institutional support for new initiatives also helped support eBird's innovative approach while advancing the organization's mission, tying project sustainability to several important contextual characteristics. Diversifying revenue streams to include multiple sources with different levels of vulnerability to external conditions clearly benefited the project, but realizing these benefits required expertise, infrastructure, and strong partner relationships. The institutional environment surrounding eBird helps support grantwriting, endowment management, contract negotiation, and access to individuals with the necessary domain and technical skills.

6.4 Institutional Influences

In addition to its organizational context, institutional influences affect eBird’s organizing processes primarily through partnerships. The broader impacts of eBird’s relationship to other members of its organizational field include third-party adoption of eBird as an infrastructure to support their own projects and collaborations. These themes are explored in this section, which expands upon the description of eBird’s organizational influences to include the broader set of relationships that have an ongoing influence on project development.

6.4.1 Institutions and Partnerships

Most of the members of eBird’s organizational field to which interviewees referred are domain-specific: Audubon, Bird Studies Canada, BirdLife International, several Bird Observatories, federal agencies such as the U.S. Fish and Wildlife Service and U.S. Geological Survey, ornithological societies, and other bird monitoring projects and programs. Several related projects are also operated by other departments in the Cornell Lab of Ornithology, including the Great Backyard Bird Count and Project FeederWatch. While these projects may appear to be potential competitors to eBird, they have different participation structures and typically target different groups of contributors, with eBirders being the most advanced and “avid” birders. The more extended network of institutions mentioned by interviewees included governmental authorities and numerous small conservation organizations. Most of these are only distantly linked to eBird, but nonetheless represent potential sources of new contributors or data users even without formalized organizational partnerships.

Through its success with eBird and subsequent projects, the Information Science department has accumulated a wide variety of collaborative partners that support its activities. These partners range from localized groups that deployed an eBird portal—putting their own branding and content around eBird’s core data submission and management functionality—

to high-performance computing groups, such as TeraGrid and Oakridge National Research Labs that provided access to grid computing resources for data-intensive analysis and modeling. Collaboration with more than a half dozen university computer science departments helped the project develop new tools, algorithms, and ways of analyzing the data with novel data mining techniques. Cross-disciplinary partnerships have revealed a variety of other areas in which the eBird data provide a valuable resource to address scientific interests. Both eBird organizers and their colleagues are involved in data mining research, developing new statistical methods, analysis of user behavior, and high performance computing. Although the project was not designed with these purposes in mind, the nature of the data that are generated, and the volume of the data, have met a broader range of scientific interests than its founders could have foreseen. These partnerships also bolstered project sustainability by providing access to new sources of funding for addressing an increasing variety of scientific interests.

Often referred to as partners, these members of eBird's organizational field sometimes provide additional resources, which are presumably mutually advantageous arrangements. In some cases, eBird was a resource and tool that other organizations use to support their organizing processes, while eBird has interacted and partnered with numerous organizations in the process of organizing its own activities.

This was most evident with eBird portals. External groups can commission an eBird portal branded for their own uses, with service provided by the Lab. Each of these arrangements represents an interorganizational partnership that brings the project in closer contact with other members of the organizational field. They represent both a revenue stream and a long-term formal relationship between eBird and an external organization. They serve the eBird project goals and Lab mission by leveraging existing social structures to increase exposure and adoption of eBird. In addition, these partnerships demonstrate that eBird is

providing valued infrastructural services to other organizations.

[They] basically came to us and said, we have the need for these tools and we have the community of people that wants to use them and we said, okay, here you go, take it. And they develop their own data quality filters, and basically we manage the hardware side of things and make sure that everything runs smoothly, but our goal is to just enable whatever group worldwide that wants to do that with these tools. (Pterodroma, 18230–18662)

These partnerships were clearly a mutually beneficial arrangement. The portals provided eBird with access to a much broader global networks of organizations, societies, clubs, and agencies that are concerned with bird conservation and research. In addition to the potential benefits for contributor recruitment and expanded audiences of data users, the portals can also lead to greater institutionalization of eBird monitoring protocols, which implicitly advances scientific data sharing.

Partnerships can also influence aspects of the scientific research; for example, Audubon coordinates with the U.S. Geological Survey to ensure that analyses of the annual Christmas Bird Count and Breeding Bird Survey are complementary (Columba, 4168–4533). Likewise, when the Deep Horizon oil spill flooded the Gulf of Mexico with toxins, the eBird team quickly implemented a new interface for reporting oiled birds, with a slightly modified protocol developed through partnership with National Audubon. In turn, Audubon worked with the U.S. Fish and Wildlife Service to coordinate use of the eBird data to evaluate the effectiveness of the strategies employed to mitigate the damage from the oil spill. Another partner project organizer noted that if their collaborators, state and federal agencies and conservation organizations, were to adopt eBird, it would further simplify data sharing among conservation partners (Circus, 40148–40762).

Providing a type of infrastructure service has been one of the strategies that supports the eBird project's sustainability. Technical staff identified related implications for technology development, highlighting a specific need:

...to be able to adapt faster to different situations. The [Deep Horizon] Gulf oil spill is a good example of that. There was a pretty immediate need to be able to record data about dead birds and oiled birds. Right now the way the UI is built doesn't give us a lot of flexibility that way. So that's one of the big things that we're working on right now...the code revamp, to be able to build that quicker, but also keep it easier to use. (Platalea, 6003–7430)

He went on to explain that a custom interface was required for each new protocol, and in planning system improvements, the eBird team reviewed similar past projects to identify common needs for rapid reaction situations. This process demonstrates the strong influence that partnerships can have on the technology itself. The influence of partnerships also arises with new features developed for eBird portals and then deployed more widely (Wood et al., 2011), which is further evidence of the infrastructural role that eBird increasingly plays within its broader organizational field.

A completely different set of sociocultural considerations come into play with globalization of the project, development of a cross-cultural contributor base, and development of international partnerships. An international portal organizer described the cultural differences in the birding community that he experienced in moving from Europe to South America:

For me, it's not possible to go in the fields without taking notes, and without sensing, counting birds, or this kind of thing. Even common birds. But this culture is not here. This culture is not in Chile. And we will need probably five, ten years to have most of the birders doing it. (Diomedea, 13700–14878)

While he was optimistic about the timeframe required to effect a cultural shift, active evangelism and effective communication, discussed later, produced good results in Chile.

In addition to the need to develop a local culture of data contribution, Diomedea also noted substantive differences between North American and European birding culture: “eBird give [*sic*] you the raw data free on the web. All the data, all the information is free on the web. In France, it's completely different,” (Diomedea, 33526–33716). He attributed the French reluctance to contribute data to eBird to a complex institutional configuration, in

which contributors provide their observations as a membership fee to local conservation organizations. These organizations then sell the data for environmental assessments, providing the groups with funds for conservation actions. Contributing the data to eBird instead (or in addition) would therefore undermine the sustainability of bird conservation organizations in France. This is a challenge with no easy solution. Despite these cross-cultural constraints, the broader institutional impacts of eBird continue to grow, as the following section discusses.

6.4.2 Broader Institutional Impacts

A few of the ways that eBird has made a broader impact include being used to influence land management and policy decision-making, such as disaster planning for chemical spills in waterways. These policy and land management applications of eBird data have also become more viable because of the increasing scale of the data: “it’s just getting to the point where we are going to see more and more information come out that will help drive policy and decision-making” (Columba, 10685–11427). The expectation and use of these data as a tool for decision-making stands in testament to its perceived value and quality. eBird is swiftly becoming the best available data set for these purposes, in addition to its contributions to scientific knowledge production. Data from eBird have been used to communicate with policymakers in addition to direct decision support. In 2011, the annual State of the Birds report to the Secretary of the Department of the Interior, which focuses on the status of bird life on public lands, was based on eBird data and signaled an increasing level of trust in and authoritativeness of the data.

These uses have also made eBird a resource that can have a broader impact in facilitating collaboration by enabling new partnership projects that were previously more difficult to organize.

We are in conversations with various groups about how we might help them get the volunteers, and put in a scientifically rigorous survey that will allow us to get

some of that information back out. So it's pretty exciting...what we're able to do because we have a tool like eBird that we can use in these situations. (Columba, 11712–14428)

This particular effect of the technology demonstrates again that eBird is becoming a form of infrastructure for formal organizational partnerships in addition to localized efforts.

Functionality that helps satisfy the curiosity of data contributors is not substantially different from that which satisfies the needs of data users who access eBird data to fulfill job duties. One such user discussed the way she used eBird for conservation work, mentioning that eBird saved substantial effort in compiling annual reports because there was no longer a need to contact individual volunteers one by one to request data from them. She summarized by saying, “I know that it doesn't do a perfect job, but so far I'm really satisfied with it,” (Circus, 11312–11421). Circus further clarified that most of her uses for the data obtained from eBird reports are so specific to her own work processes that it would be unreasonable to expect anything more with respect to data outputs.

On the other hand, there are some challenges associated with appropriating eBird for existing monitoring efforts. For one partner organization project leader, using eBird to organize volunteers for location-based projects required a number of subtle but meaningful adaptations. These included the use of a shared account, reducing data entry by recording a series of stationary counts as one traveling count, and recording a variety of additional data points in the notes section of the data entry interface. The reasons for these adaptations was that “[eBird] has certain ways of entering the information. ...I'll just tell volunteers to put it in this way, or put it in that way, and I know what that means. And if there's any extra information that you gathered, to shove it all in the notes section” (Circus, 25257–26049). eBird was preferred to an existing purpose-built database that would fully accommodate these details because of its ease of use for volunteer data entry. Unfortunately, because of

these adaptations, the comparability of these data to those generated by other eBirders is questionable, and these differences are invisible to others who use eBird data. The situation discussed here suggests that finding ways to accommodate customized uses of eBird may be preferable to supporting appropriation of the technology that could dilute data quality.

A project leader mentioned another example of using eBird for a separate research project, discussing a high school student's use of eBird:

He developed a separate data entry mechanism using Google Docs that then reformatted the data so that it could be uploaded into eBird. He was able to develop a project that in many ways was more specific and rigorous than what could be gathered by eBird. His genius was to demonstrate how someone could develop far more rigorous protocols for eBird to answer specific scientific questions, and then still allow these data to be aggregated for broad scale questions eBird hopes to answer. (Pinicola)

This project, led by a resourceful student, shows another type of contribution with broader impacts on scientific research.

6.5 Designing Technologies for Participatory Science

As a partner project organizer put it, “as technology has come along the birders have adapted that technology in new ways” (Columba, 41834–42184). eBird has been successful in matching recreational birders' personal interests to researchers' interests; another organizer commented that “it serves birders well, and it serves the scientists well” (Elanoides, 33829–33892). This alignment of interests was an important element in eBird's adoption by both birders and scientists. Several lifelong birders described the evolution of technology use in the birding community, clearly and consistently identifying eBird as not only the current state of the art in birding technologies, but also as the foundational infrastructure for future innovations.

eBird is not designed to answer a specific scientific research question, but to provide

data that can answer numerous research questions: “eBird is a surveillance, monitoring project. We didn’t develop eBird to test a specific experiment” (Dendroica, 5166–5825). At the same time, the BirdSource project that developed into eBird was initiated to address the hypothesis “that bird watcher observations can have significant value in studying the patterns and trends, and can be used for the conservation of birds” (Dendroica, 1994–2867), which has subsequently been demonstrated to be true. This section describes the importance of aligning scientific and personal interests, and the resulting impacts on technology design, science processes, and scientific outcomes for eBird.

6.5.1 Aligning Scientific and Personal Interests

The initial assumption that only scientists are interested in these data proved wrong; birders are also interested in the data for a wide range of purposes: “Initially eBird was developed by scientists for science-minded birders. They thought well, people wouldn’t be interested in such things, but they really are” (Pterodroma, 4539–4713). The primary challenge was to establish participation processes that can generate data appropriate for addressing scientific interests, while also being suitable for recreational birders to carry out, and then translating those processes into a usable technology. The eBird project leaders worked closely with statisticians, modelers, and biologists as well as the broader birder community to keep these interests in balance as the technology continued to evolve.

eBird is a technology with accompanying participation protocols based on the scientific needs for the data as well as the norms for participation in the birding community. The combination of community practices and scientific processes is somewhat unique, as citizen science participation protocols tend to be more science-centric and less community-oriented; this point is discussed later with respect to designing a project for a community of practice. The protocols for participation are “loose” and opportunistic from a scientific perspective, in the sense that they permit contributors to participate in a number of different ways, where

and when they choose, contributing data with varying levels of specificity and scientific value.

The basic form of participation is to submit a checklist of birds, preferably with counts, and preferably including all birds observed during a given period of time in a specific place, and over a specific area or distance. However, eBird also accepts “incidental” observations of one-off sightings of species without any metadata, which is important to supporting the full range of data that birders are interested in keeping. Supporting multiple variations on a protocol is unusual in citizen science, particularly within the boundaries of a single project, but is self-evidently sensible in eBird because the protocols are built on existing community practices which include multiple modes of observation.

As a result, the observation protocols for eBird have been fairly stable since the project began, and only a few minor changes have been made to protocols to improve data quality. Notably, there is a distinction between the observation protocols and the choice of protocols that contributors make. For example, through collaboration with biologists and statisticians using the data, the eBird team learned that the data most valuable for statistical analysis are repeat observations taken at the same location over time, as would be standard practice for many conventional scientific research protocols. In order to improve participation in collecting the most scientifically valuable data, new website features (e.g., Yard Lists) were developed to reward contributors for greater frequency of repeat observations in fixed locations. This did not represent an actual protocol change but rather a change in the way that participation was framed, making it more congruent with the personal interests of the birding community. While the process of observation remained the same, birders were encouraged to contribute data more repeatedly for specific locations, with a net result of increasing the value of the data through accumulation. The alignment of scientific and personal interests has also had meaningful impacts on system design, discussed next.

6.5.2 Impacts on Technology Design

eBird's unique functionality is considered highly innovative by members of the birding community, as reflected in the interviews and listserv postings. During formal and informal interviews, the system was frequently identified by other citizen science organizers as the most sophisticated tools they had encountered to support observation-based citizen science projects. Although this distinction is often attributed to the fact that eBird has had substantially greater funding and more time to mature than similarly structured projects, it does not diminish the appreciable accomplishments of the project team. The project leaders consistently put the highest priority on developing functionality not available elsewhere. As an example, visualizations—particularly graphs and maps—are among these essential features because, as one organizer explained:

People like maps, and they've always liked maps, and if you can animate the maps they even like them more. So we spent a lot of effort and got recognized for our ability to visualize these kinds of patterns, spatial and temporal patterns. (Dendroica, 18417–18713)

It is difficult for a non-birder to appreciate the innovativeness and resultant delight these features inspire in avid birders. One of the project organizers, self-described as a casual birder, emphasized the novelty of eBird's range maps, released in the summer of 2010, calling them “unprecedented. Nobody has ever made range maps built on hard data, this much observational data. ...There's a lot of guesswork involved in making range maps, and here, just look at the detail on those. It's really exciting” (Fusca, 40944–41286). The birders among the interviewees unanimously expressed great enthusiasm for these unique features, calling them “incredible,” “fun,” and “innovative.” Likewise, comments on the eBird TechTalk Google Group reflect similar enthusiasm for each new feature, as well as continual demand for more.

The eBirder's excitement over eBird features can be directly connected to the strategic choice to hire project leaders who are established members of the birding community; as one project leader put it, "when I say *we*, I mean the birding community" (Pterodroma, 1049–1093). This choice, combined with redesigning eBird to satisfy the birding community's known interests, has proven highly successful. While this staffing strategy may appear obvious and logical, it was not how the project was initially arranged and was a pattern that was repeated across cases; later its implications will be discussed further.

Every member of the eBird technical staff repeatedly mentioned that the project leaders are one of their greatest assets due to their deep engagement with the birding community. The question of how new features are evaluated in eBird's development process was met with this explanation:

A lot of it is just Pinicola, Pterodroma and Stercorarius, and our programmers who are rabid birders, it comes from those guys being their own customers, and just trying to enter data through the site, and they know where they get frustrated. ...It's like you build a tool for yourself and then solve your own needs, and as long as you are within that target demographic, you're actually probably solving a lot of the same problems that other people are having. I couldn't do that, I'm not a rabid birder. (Platalea, 16969–17573)

Notably, the eBird team are akin to open source software developers in this respect, as they are user-experts who can serve their own interests by helping to improve the system. The similarity to open source also extends to enthusiastic users who have the technical skills to provide additional value, as several individuals have shared their Google Gadgets and techniques for using data export tools with Excel to generate additional outputs of interests to birders.

The degree to which eBird's design shifted from a scientifically-minded tool to a community-oriented technology that simultaneously supports scientific interests is best demonstrated by its specialized taxonomy: "the eBird taxonomy is a morphed version of that [Clements]

taxonomy...it's been changed to serve the birdwatching community," (Fusca, 7466–7636). Taxonomies are very important in the biological sciences, but in order to improve both the user experience and scientific outcomes of the project, eBird's taxonomy has been customized. The specialized taxonomy includes less specific taxonomic categories for birds that are difficult to identify in the field. For example, *Empidonax* flycatchers are very difficult to distinguish in the field even for world-reknowned expert birders (Kaufman, 1990), so they can be reported as a "spuh" (*Empidonax sp.*) meaning an unspecified species of flycatcher.

The use of "spuhs" is an eBird-specific convention to support data quality: "We are not going to do any great science with the sparrow spuh, but we are going to keep someone from just pigeonholing a bird because they think that's what it is but they're not really sure." (Stercorarius, 6355–6670). This is an important combination of technology and participation design that acknowledges the uncertainties of field observation and variability in observer skill by accommodating multiple levels of detail in species identification, much as the protocols and data entry allow flexibility with respect to counting birds versus reporting simple presence and absence. The customized taxonomy also allows birders to track subspecies and hybrid species that are of interest to recreational birders but are otherwise generally regarded as uninteresting by professional ornithologists, demonstrating further alignment with personal interests. In many ways, eBird is an ideal example of designing both the scientific data collection and the supporting technologies in order to meet community needs and align with community practices. The careful alignment of interests with system features has generated an enormous data set, the characteristics of which influence scientific research processes, as seen in the following section.

6.5.3 Impacts on Science Processes

The link between contributions and science is quite clear; the collection of data is a critical step to any research process. In particular, the scale and scope of the contributions

was mentioned several times as a strong influence on scientific processes. Statistical modeling and high-performance computing were discussed as solutions for working with very large data sets that exhibit such wide variability.

The availability of such a large dataset has led to changes in the research approaches used for working with it. In reference to the shift toward data-intensive analyses using high-performance computing, a researcher who uses the data said, “if the technology makes new things available, you change your focus to exploit it” (Passerina, 2822–3677). In addition to the perennial scientific concerns about the geographic and temporal biases of the data, missing data, and size and scale of the data set, he also noted the fundamental challenge of the scientific research itself:

We’re trying to study bird migrations. Forget the variability in the data: we have a network of people out there, and they followed the perfect protocol and told us exactly where the birds were and when they were, recorded everything perfectly. Even then there are some challenges just to understand, to model the dynamics of the birds themselves. They really vary a lot, it’s a very rich phenomenon. (Passerina, 12638–13276)

With respect to technical infrastructure that supports the accumulation of such a large data set, issues of scalability and extensibility were also mentioned in relation to sustainability. Supporting ongoing project growth and development requires the technical infrastructure to be scalable and extensible to meet new and increasing demands. Speaking of the AKN and eBird Reference Dataset, one of the technical staff speculated that the current hardware for the data repository “would be probably overwhelmed if we didn’t have the static data sets bleeding off some of the usage” (Meleagris, 37894–38071). On the back end for eBird, additional infrastructure was added as the project grew in scale: “we beefed up our database server...and that really made a big difference. We were getting seriously disk-bound before we did that” (Fusca, 20229–20427). These challenges highlight an important consideration for citizen science projects that are reliant on technologies to support large numbers of con-

tributors; success in developing a contributor base can also lead to rising costs for project maintenance and sustainability as project scale increases.

The tradeoff for such a large volume of data lies in quality: “overall it’s a balance between getting as much data as we can get, and it’s also looking for the best data we can get” (Passerina, 30315–30897). As a researcher, Passerina noted that the variability of observer expertise and by extension, data quality, is his primary concern. The project was designed with “opportunistic” protocols, and has been met with similarly opportunistic research approaches, such as nonparametric and semi-parametric statistical modeling. While the opportunistic nature of sampling and leniency of protocols support contribution volume over quality, the review processes have helped balance concerns over data validity.

Supporting research with large-scale data sets requires a carefully structured and essentially unchanging participation protocol so that longitudinal observations are comparable. The scientific outcomes that have emerged from the project range from avian epidemiology to data mining algorithm development, demonstrating that in addition to serving the ornithological research community, these data have been valuable to research in other disciplines as well. This means that the scientific knowledge outputs have been more diverse than originally predicted, with the broader impacts of the project including science-based conservation and management decision-making, as previously discussed. The impacts on traditional scientific knowledge production are the focus of the following discussion.

6.5.4 Impacts on Scientific Outcomes

The design of the participation and supporting technologies has clearly influenced the scientific outcomes of the project. One of eBird’s initial goals, as part of the Citizen Science Online NSF grant, was to test the hypothesis that recreational birders could contribute scientifically valuable data. The participation protocol is opportunistic and general; accordingly, “there’s always been a lot of discussion on whether that data can be useful for anything”

(Dendroica, 5320–5628). The data have in fact proven valuable for scientific research, yielding at least 40 (known) scholarly articles between 2003 and 2011 that used the eBird data or focused on eBird itself. The data are also valuable for conservation purposes, such as land management and environmental policy decision support, which are generally approached through scientifically rigorous data collection and analysis. As a partner organization leader noted, “Cornell brings a lot of good scientific rigor to what they do, we bring a conservation focus...based upon good scientific rigor, and so it’s a good partnership for figuring out how to move things [conservation] forward” (Columba, 8098–8345). Therefore, although some of the research motivations for eBird data use are not scientific *per se*, they make use of scientific approaches to using the data.

In addition to publications in the scientific journals, eBird is the mainstay of the AKN. The eBird data is also the foundation of a value-added data product, the eBird Reference Dataset, that provides ready-to-use observation data that Lab staff have packaged with numerous covariates by location (e.g., weather, ground cover, human population, etc.) Project technical staff reported that the additional investment in data management seemed to be paying off for researchers (Meleagris, 12481–12726). These data products are another substantive contribution to scientific knowledge production.

In some cases, the eBird data are better than anything found in the published literature for both academic and popular audiences. For example, “what is shown by eBird in Chile is not shown by anything else, even publication [*sic*]. We have a few maps for some species, which are much better than any maps ever published, ever, for some species” (Diomedea, 9217–9715). Diomedea emphasized that these data had been collected in only 18 months after the localized portal had launched, an impressively fast pace for producing detailed range maps for species whose locations were previously poorly documented. This type of basic descriptive information about species distribution can become a foundation for further

study.

While all data contributed to eBird have the potential to be useful for any of these purposes, some data are not used for research. The main issues lie with biases in the data and the availability of metadata specifying the effort expended on collecting observations. With respect to biases in the data in some geographical regions, observations are simply too sparse for statistical modeling. The general absence of data for North and South Dakota, for example, was specifically mentioned by several interviewees, and is directly related to sparse human population in those states. The way that feedback that is provided by the system encourages more detailed data collection, which is seen as particularly valuable for statistical modeling, because greater geographic resolution supports a wider range of analyses:

People really do want to understand, not just what the species is doing all the way across country, but at really fine level of detail, too. And the finer detail that you look at in the information, you start to realize that at finer [geographic] resolution, the data is sparser and sparser and sparser. The birders love doing this, so there's this drive to get more and more. (Passerina, 17242–17745)

The eBird team has taken several approaches to addressing the issue of geographic bias at multiple stages of the research process, including targeted outreach for participant recruitment and advanced statistical modeling.

Another type of data that is rarely used for scientific research are observations submitted without metadata about the time and duration of observation: “the one [protocol] that is least used in analysis is the incidental counts, because basically it bears no effort information, and so it's really hard to statistically justify your assumptions when you aggregate that together” (Ceryle, 24508–24771). These observations are nonetheless critically important to adoption of the system by birders. If eBird did not permit birders to record the incidental observations of “life birds,” the individual first-time sightings of a species that a person has not previously encountered (e.g., a scissor-tailed flycatcher spotted briefly through the

window of a moving car), it would have had more trouble achieving critical mass with respect to adoption by expert birders, as was seen early on in the project's history.

The openness to data that do not fit the scientific ideal is one of the secrets to eBird's success: "eBird really works...on whatever level. If you just want to do one list from your backyard in a day, or once a month when you go somewhere interesting, it really works on all those levels, every contribution helps" (Stercorarius, 17630–17860). Some contributors have also entered historical data up to a century old (the earliest data eBird can accept start with the year 1900). For example, an organizer spoke of two Audubon volunteers in different regions who "saw the value" of contributing historic data to eBird, and independently took it upon themselves to enter all the accumulated data for their favorite Important Bird Areas². These historical data are considered very valuable, but are currently too sparse to rely upon for policy and management decision-making. The development of a system that is pleasing to the birding community, discussed next, has motivated increasing contributions that make decision-making support an increasingly valuable broader impact of eBird.

6.6 Designing Technologies for a Community of Practice

Birders love birds. One of the eBird staff put it particularly well:

This data that I take care of is among the most loved data, I think, anywhere. People have spent just millions of hours of accumulating it. It's really irreplaceable. And they're passionate about it. ... People are doing this because they love it. (Fusca, 29785–30339)

For eBirders, this love of birds is reinforced and shared through participation in eBird. The user base is a community of practice that centers around the observation of birds, a practice fueled primarily by their common interest.

²An Important Bird Area is a natural habitat of global importance for bird conservation. The program was developed by BirdLife International and is administered in the U.S. by National Audubon; the globally-recognized IBA designation leads to habitat protection under national legislation.

This section describes birders' shared practices, skills, biographies, and personal interests, and the characteristics of the larger birding community that the eBird organizers have targeted as the primary audience for project participation. It discusses the importance of both direct and indirect feedback to supporting ongoing participation and improvements in data quality. The community-based recruitment strategies that the eBird project leaders have implemented further demonstrate the value of community-oriented technology design.

6.6.1 Birders

As the previous section indicated, designing eBird's technologies to support participatory science was built on the foundation of a deep knowledge of a community of practice: birders. Birding is a multi-million dollar leisure industry and the fastest-growing hobby in the U.S. (Weidensaul, 2007). Birders constitute a distinctive subculture, ranging from casual bird watchers to avid listers who maintain lifelong daily records of observations, some accumulating a huge volume of observational data spanning decades. The birding community has strong norms around reporting sightings, mentoring novice birders, and rewarding contributions with public acknowledgement. Local bird clubs provide in-person contact and socialization, while bird festivals and competitions provide opportunities to engage with the broader community. Among serious birders, information is the foundation for reputation, with community status established through lists of bird sightings at varying geographic and temporal scales.

eBird serves as a tool for the birding community, which long predated the technology. As such, there are many existing social organizations, opportunities, and venues in which birders interact. eBird's direct support for non-task activities such as socialization and mentoring is minimal, likely because creating an online social network was not required to support project activity due to the extensive existing social network among birders. The connection between eBird and the birding community was described as "an incredible social network, in terms of

collecting information about bird distribution all over the country, and tying people together in a way that they have never been tied together before” (Columba, 39194–39478).

Birding is a challenging pastime. Skills are gained through experience and practice, and lifelong birders described a learning process not unlike an apprenticeship.

You really need to go out in the field with more experienced birders who can give you some clues as to why that one’s a little different. And some of them [species] are quite variable, so song sparrows look quite different here than they do in different parts of the country, and they have different races. So the more experience you have, the better you can do. You just can’t substitute looking at a book or website for 10 years experience. (Elanoides, 10665–11100)

The acquisition of these skills is a nontrivial effort that requires time and patience, and although eBird can be used by birders of any skill level, it is less likely that beginning birders would choose to start with eBird. The ability to identify more species than the average person on the street is a practical prerequisite for eBird use, which impacts the project’s participation processes and sustainability, as well as its accessibility to non-birders. Project evaluators, however, highlighted the notion that required skill or expertise can be leveraged for inclusivity rather than exclusivity (Allison-Bunnell & Thompson, 2007), as inclusion has always been the intent behind eBird’s project design. In actuality, a variety of skill sets and expertise are needed to fuel a project like eBird, and contributors bring not only birding expertise but also local knowledge.

An example of the importance of these skills is the role of the regional reviewers, skilled birders recruited by the project leaders to assist in data quality management. When their personal networks fail to yield the needed contacts for a given geographic region, they turn to eBird and mine its data to identify potential candidates.

You can pretty much tell from someone’s eBird signature what their experience level is. The kinds of things they write, the kinds of things they report regularly, or don’t report regularly. And usually by that point, if we can identify someone there [a geographic region] then we probably know someone who knows that person, at

least on some level. We say, do you know this person? (Stercorarius, 22301–22672)

Even after identifying a potential reviewer based on their apparent skill, however, the project leaders check with a local contact for further verification of the individual's abilities.

In addition to birding and technology skills, local knowledge and other individual intellectual resources are important to ensuring project success. For example, domain expertise is not required for the hotspot editors who maintain the lists of local publicly accessible birding locations:

You don't have to be an ace birder, so much as you just need to know your local geography. ... So that's kind of a way to engage intermediate, or even beginner birders, that are enthusiastic about it but might not be qualified to judge someone's Baird's sandpiper. (Stercorarius, 24911–25203)

The diverse skills and knowledge required for the roles of regional reviewer and hotspot editor provide an opportunity for role advancement, a known best practice in job design (Ilgen & Hollenbeck, 1991) that extends to volunteers as well as paid employees. These meta-contribution roles also provide recognition for leaders in the local birding communities, acknowledging their expertise.

The biographical characteristics of birders were mentioned for two distinct groups, both the project organizers (interviewees), and the participants with whom they interact. When the eBird project organizers (staff and related individuals in leadership roles) talked about their background with the project, most referenced their educational background and connections to birding. The organizers who were interviewed were a diverse group, with education and experience in areas such as forestry, biology, computer science, natural resources, ecology, statistics, studio art, and neurobiology. As previously discussed, having team members who are members of the birding community was viewed as valuable by those organizers who claimed to know little about birds, and was considered particularly important as a characteristic of the project leaders. According to one interviewee,

Those three guys are also hard-core birders, and they are also users of the site. And they are probably our three most challenging users that you could possibly have. They have very little patience for things not working properly...and so they are actually my most valuable resource. (Platalea, 15221–15594)

All of those who self-identified as being birders mentioned that they have been involved with birding since their teen years, demonstrating a long-term personal interest and implying a wealth of accumulated skill and knowledge about birds and birding.

Although there was relatively little discussion of the biographical characteristics of eBird contributors, factors that came up in conversation were consistent. Most organizers connected age or generational membership to technology self-efficacy (Compeau & Higgins, 1995), or being “computer savvy.” Most people asserted that older adults are generally less skilled in computer use and said that younger birders have little trouble using newer technologies.

Several noted that the majority of the participants with whom they interact are age 40 and over, suggesting that many participants may not be particularly comfortable using computers and related technologies. The perceived influence of age on participation relates to another biographical characteristic of some individuals, relevant because of the value of data from lifelong historical records. This trait was described by one project leader as “people that have that gene...for recording information” (Stercorarius, 34902–34996). Such individuals were also mentioned by other citizen science organizers interviewed in this study as being an ideal participant, in part because these individuals enjoy the activities of both collecting observations and also making records of them, a primary task in most citizen science projects that is often considered dull.

The eBird case, however, highlights a situation where an innate interest in or tendency toward record-keeping can also be a barrier to participation. Organizers mentioned more than once that if a birder is already comfortable using notebooks or desktop software, it is

harder to convert them to using eBird, despite the availability of bulk import tools:

The simplest ones to convert are young birders who are computer and tech savvy and it's like second nature to them, they are like 'yeah, this is a no-brainer, why wouldn't I do it?' The most difficult ones are older birders who have been doing things their own way for 40 years and have their own system that may not be necessarily computer-driven. It might be a series of notebooks or something like that, and this is a major hurdle for them, to get a computer and get online, digitize all this stuff, and change the way they're doing things. (Pterodroma, 19672–20226)

One of the project leaders also noted that while an increasing proportion of the birder population is comfortable with computing technologies, it is still important to reach out to the senior birders because they serve as mentors to younger and novice birders, teaching them good field observation habits and communicating the value of contributing their observations to science.

The motivations to contribute data to citizen science is a topic of much interest among researchers. The norms of data sharing in the birding community provide social motivations for contribution that complement birders' personal interests. In a final report for the Citizen Science Online grant that provided early support for eBird, evaluators found that in addition to helping manage their personal bird data, contributors chose to use eBird because the data that they collected can be shared with like-minded people. The same user survey found that the contribution to science was rated as a very important benefit of participation more often than maintaining and organizing personal bird records (which took a close second place for "core" eBirders who make substantial data contributions) or learning where birds are being seen. This finding was consistent across both eBirders and non-eBirders who use the system infrequently. The desire to share a personal passion with others can be a strong social motivation for participation, and there were several mentions of individuals spending quite a lot of time entering historic bird observations. As one project leader reflected, "there's also point for people where they say, I've been doing this for so long, what's it all going

towards, if I don't share it?" (Stercorarius, 37707–37836). Posts to the eBird TechTalk listserv directly supported this observation, mentioning the motivation of leaving a legacy of birding observations submitted to eBird as having value both today and forever.

In addition, contributors had mentioned the motivation of helping with the scientific goals of the project, not just to organizers, but also more publicly in posts to the eBird TechTalk listserv. The link between contributing data to eBird and their own personal interests is clear: by contributing data, they enable research and conservation actions that support their passion. As a Lab staff member explained, “they are happy that the scientists can use the data. Because you know if there's not any birds, what you got [*sic*] to do with your time? Play golf.” (Elanoides, 34050–34211). Communicating the relationship between participation and conservation is the subject of the next section.

6.6.2 Communication is Critical for Organizing

The care taken to design eBird's technologies to serve a variety of different audiences within the broader category of birders and assure that it is equally functional for daily and occasional contributors extends to the way the project organizers communicate with participants. Both indirect communication through system-generated feedback and results and direct communication with community members have been important tactics in building eBird's contributor base.

The eBird team have worked to support sustainability of participation by including feedback to contributors as a fundamental part of eBird's design. As the organizers noted, many citizen science projects fail to provide adequate feedback despite awareness of the importance of this type of communication. Comprehensive access to data and reporting tools is one of eBird's most distinctive features. The reporting tools were directly inspired by the known personal interests of birders, who enjoy exploring information about birds. As one of the project leaders noted,

I think the more you can make people enjoy the project and get some reward back out of it for engaging with the project, then the better off you'll be for sustaining it. And we've seen significant growth that hasn't slowed down since we turned the switch on and sort of changed the way we think about it. (Pterodroma, 39971–40737)

Communicating results and feedback to contributors was often highlighted as a critical part of the participation design. In the words of an eBird portal organizer,

They need a result. You can't just tell them that the datas will be used by scientists or whatever. They don't see scientists, they don't see the paper. So you have to show them what is done with the eBird datas, and you will have to acknowledge people also as much as possible." (Diomedea, 12806–13266)

In particular, providing instant gratification through rapid or immediate feedback was cited as a powerful motivator for ongoing contribution. This appears to be a two-way street; eBird technical staff remarked that the project leaders often share user feedback with them, which they find particularly rewarding and motivating.

eBird project organizers use several channels to communicate directly with contributors and data users. Content management functionality is embedded in the eBird website, and several members of eBird staff expressed respect for the quality of the articles that the project leaders write for the site. These articles provide recognition to exceptional contributors, tips on making difficult identifications, results from research using eBird data, and announcements of new functionality. The eBird project leaders also maintain a separate blog called "Chip Notes," an eBird Facebook fan page, and Google Groups email lists for regional reviewers and for questions about data entry and protocols. They each noted that email responses to queries from contributors, data users, and members of the broader birding community require a substantial portion of their time. By anecdotal comparison to other online communities, the eBird project leaders are exceptionally responsive to these inquiries, both by direct email and on listservs. Organizers for partner projects further promote eBird via

listservs, electronic newsletters, print magazines, and in-person presentations. In addition, the eBird team conducts periodic user surveys, directly asking contributors for feedback that influences development priorities.

Strategic communication with contributors has also yielded an improvement in the usefulness of eBird data for scientific analyses. Several eBird staff recounted the story of a campaign mounted by the project leaders to educate users on how to choose appropriate observation protocols and the associated scientific value of reporting data using the more formal effort-based protocols, rather than the “incidental” sightings, which have no effort data and are therefore not particularly useful for research. The strategic communication campaign was very successful: staff showed a graph of the frequency of checklist types in which the trends were completely reversed as a result of this intervention. One of the other applications of communication by project leaders is discussed next: community-focused communication as a recruitment tool.

6.6.3 Community-Based Participation Recruitment Strategies

For eBird, like the other citizen science projects in this study, standard volunteer management practices like creating volunteer job/task descriptions, managing risk, and direct supervision simply do not apply in a meaningful way, largely due to the nature of participation as a form of distributed work in which nearly every individual performs the same fundamental task. Rewarding contributors is built into the system in the form of access to data and visualizations, rankings on leaderboards, and for a few devoted contributors, a profile on the eBird homepage as “eBirder of the Month.” Among other volunteer management processes, retention seems to be less of a concern due to the design of eBird to appeal to birders’ personal interests; orientation and training is minimal, as it is self-guided with online materials, although outreach events and presentations often include an introduction that serves as training; and screening and selection are required only for meta-contributors.

Recruitment, however, is an ongoing area of interest, and not only in the context of globalization.

In order to achieve a critical mass which would permit the project to become an authoritative data source (and garner increasing participation), the project organizers initially focused on recruiting and supporting expert birders who need relatively little education with respect to ornithology and the merits of the scientific process (Allison-Bunnell & Thompson, 2007). The strategy for recruiting contributors in North America has since evolved from this initial focus, which involved recruiting contributors from areas where there are large, active birding communities, e.g., New York and California. To expand the contributor base and further improve the scientific value of the data, one strategy the project leaders are using to address the geographic biases of opportunistic observation is through geographically-targeted volunteer recruitment:

Our goal now is to shift from focusing on outreach in areas where we know there are a lot of birders to...trying to get a more evenly distributed sample across the landscape, by engaging groups in North Dakota or Oklahoma to try to start to fill some of those holes. (Pterodroma, 21784–22954)

Recruiting contributors to fill geographic gaps in the data promises to be a much more challenging task. The easily converted birders have already signed on, leaving a target population that is most likely less geographically centralized, less interested, less confident in their skills, or simply less aware of eBird.

Across organizations and roles, project organizers leveraged professional, scientific, and birding-specific network connections to aid the development of the project; several interviewees were recruited to work on the eBird project through network contacts. The project leaders relied on the social networks they had developed through years of interactions with the birding community, particularly for volunteer recruitment for the regional reviewers, “playing on contacts that I’ve made over my years birding, trying to talk people into helping

us” (Stercorarius, 18845–18956). Partner organization leaders also used both formal and informal networks to recruit eBird participants: “When we first started, I was contacting a lot of them, the main birders in Canada, and telling them to check it out. All of them, I think, checked it out” (Otus, 10174–10401). In each of these examples, leveraging personal contacts made through community involvement provided valuable resources to the project.

The “evangelist” role was often mentioned by those organizers who recruit data contributors, the primary human resource upon which eBird’s success depends. These organizers, who know the extent of the birding community, felt that current participation is low compared to what might be possible, and believed there is substantial human resource capacity (or “cognitive surplus” (Shirky, 2010)) that could be harnessed by extending the eBird contributor base. They remarked upon the need for paid staff, both in local organizations and at the Lab, for ongoing project coordination and communication.

Among to the communication strategies discussed in the prior section, the necessity of in-person outreach for recruitment was surprising given the reliance on technology-mediated participation. Notably, these outreach efforts are primarily about recruitment rather than education, which stands in contrast to the usual meaning assigned to outreach in some citizen science projects. Giving talks at meetings was mentioned by several interviewees as a primary outreach tool:

It takes a lot of evangelism. Between the three of us [project leaders], I’ll bet we did probably 60, maybe more, eBird talks around the country in the last fiscal year. And those vary from keynote presentations at birding festivals for hundreds and hundreds of people, to very small bird clubs...where 20 people might show up. (Pterodroma, 20406–20773)

Another eBird project leader estimated that he spent about 100 days per year traveling to attend and speak at events, run training sessions, and present to groups (Pinicola). While time and resource intensive, this form of outreach takes place within the existing structure

of the birding community and related institutions and organizations, which is a particularly powerful approach to organizing collective action (McAdam, 1999). It therefore comes as little surprise that communication was often spoken of in the same breath with community and networks.

The connection between community and networks highlights a particularly effective recruitment strategy. Through the project leaders' continual outreach efforts, other organizations are beginning to promote eBird: "a lot of those groups [state ornithological societies] are beginning to support eBird, telling their members to report on eBird" (Elanoides, 42215–42506). This is an example of tapping into "indigenous organization strength," a concept from social movements theory (McAdam, 1999). An indigenous organization is one which exists in the community, well established prior to any attempts at organizing for collective action; classic examples are churches, sports teams, and campus groups. These organizations bring four crucial resources to mobilizing collective action: members, leaders, communication networks, and established structures of interpersonal rewards that motivate participation and solve the "free rider problem" (the question of how to prevent consumption of resources without contribution) (Olson, 1965). Birding groups and bird conservation organizations fit this description quite well, and cultivating relationships with them has yielded similar benefits for eBird.

In addition to North American outreach efforts, eBird's successes to date in globalization have been based on the strength of indigenous organizations. Leveraging these relationships remains the primary strategy for expanding the project's reach more globally. In the future, recruiting new contributors will likely increasingly rely on these networks as a way to reach out to a wider audience of bird enthusiasts. In the more informal context of birding listservs, checklists emailed from eBird include a footer denoting the source of the observations; interviewees claimed that in some areas, most of the posts to listservs are generated

by eBird, potentially exerting a subtle social pressure on listserv members who do not yet use eBird. An eBird user made a listserv posting that substantiated the claim that users' communication on birding listservs can convince others to participate, as he had received many messages from more experienced birders informing him that they had adopted eBird on his recommendation. Endorsement by fellow birders is not the only incentive for adopting eBird, which is the topic of the following section.

6.7 Participation Incentives and Transformative Experiences

eBird is explicitly designed to reward contributors. The direct benefits of participation include list management, rare bird alerts, and other tools to learn more about birds. Most respondents to eBird user surveys conducted by the project organizers report that maintaining their birding records, keeping an eye on what other bird watchers report, and finding general information on bird distribution and abundance are their primary activities on eBird, with tracking personal birding records being the most frequently cited reason to use the system. This section discusses the ways that the eBird system design not only incentivizes participation, but also promotes increased commitment, changes in birders' behavior, and individual development.

6.7.1 Incentivizing Participation

The initial technology design for eBird relied on altruistic intentions to support “birding for science.” After redesigning the tools to provide functionality that interests birders, project performance increased dramatically. The practice of rewarding contributions with appealing functionality is one of the core design principles for the project. Project staff were quick to dispel the notion that relying on altruism could sustain ongoing participation, saying that “people will be excited about it for a while, but finding people to participate over the long haul, if you're just counting on altruism, I think it's not going to fly” (Fusca,

39241–39925). Viewing the relationship between the project organizers and contributors as an exchange of data for tools has perpetuated a strong user-centered design ethos.

The underlying strategy for eBird’s technology design was best summarized as, “let’s give them the tools to do what they want, and they’ll give us all of their data” (Passerina, 26897–27937). At the same time that the technology was substantially redesigned in each subsequent version of eBird, there was essentially no change to the core observation protocols, except to update labeling and descriptions of the participation protocols for clarity. This is unique in the world of observation-based citizen science projects; nearly every project reports a revision cycle with respect to participation protocols. But not eBird.

The main point of differentiation from other monitoring projects is that eBird’s observation protocols were based on the existing, long-standing practices of the birding community (Wood et al., 2011). In essence, birders were already doing the work that eBird asks of them, but were not previously recording and reporting the data to a centralized database. The eBird project does request additional information that is not usually recorded in the field, but for the minimal level of contribution it requires only a relatively minor modification of the usual birding practices, depending on the birder. Over time and with gentle prompting from the project leaders, the less valuable incidental observations have been supplanted with effort-based checklists, as previously discussed. This shift does not represent a change in protocols, but rather a change in the choices of participants as to which protocols to use and the degree of detail to contribute.

The approach taken by eBird organizers represents a different tactic from most design models for citizen science. The typical project design approach in citizen science is to identify a scientific problem, and create a way for people to participate in scientific data collection or data processing tasks. When eBird was designed, organizers were able to take the opposite approach. The target participant community’s existing practices were used as the basis

for designing participation tasks that were minor extensions of the existing traditions in the community. Requiring only moderate, gradual changes from the way a lifelong pastime has been practiced means it is easier for contributors to incorporate project participation into their everyday routines. This is a more robust design strategy for supporting ongoing participation than the contrasting approach, where contributors are asked to undertake completely new tasks with no real relationship to existing practices or habits.

Like the GSP and Mountain Watch organizers, the eBird project leaders highlighted the tension of design tradeoffs between participation and science. “We always walk this line at eBird between usability and utility. We want people to collect better and more valuable observations, but the more you ask people to do, the fewer people will actually do it” (Pterodroma, 6298–6840). The strategy that project organizers took in response to this persistent challenge was to accept data collected according to protocols with variable rigor, allowing contributors to choose the level of detail that they record.

At the same time, they clearly understand how to motivate birders to choose a more rigorous participation task:

What we found is that the more people that we can get to survey a location multiple times, the more detailed the data, then the more valuable the data is for analysis. So last year we tried this site survey concept, and that didn’t work very well. ... So this year, we’re going to turn it into a game, where you can keep track of how many birds you see in your yard, compare that with others, and get your name on the list. (Dendroica, 20565–21140)

While the new functionality is fun for many eBirders, it was inspired by a scientific data need. In contrast, both scientific interests and personal interests have also been served by feature development prompted by user requests. eBird project leaders regularly invite feedback from eBirders on development goals, and this dialogue between the organizers and community helps ensure that resources are targeted toward the development goals that will serve the mutual interests of the project organizers, data users, and data contributors. One

such goal is making data entry easy and rewarding, the focus of the next section.

6.7.2 Making Data Entry Worthwhile

While most of the eBird participation process closely mirrors community practices, the step that represents a substantive and unwelcome addition to most birders' existing habits is data entry. Data entry is considered “a hurdle, there's no question about it. And the easier you can make that process, the better,” (Pterodroma, 6298–6840). As reported by both interviewees in this study and organizers in the broader practitioner community, the distaste for data entry is a universal issue across citizen science projects that involve data entry as a separate step from observation. eBird's answer to the data entry issue was to ensure that participation is immediately rewarded:

There's only so many people out there that will spend their time sitting behind a computer doing data entry because they think it's good for the birds. But there's a whole lot more that will spend their time sitting behind a computer entering data if they can then get something out of it that they find valuable. (Pterodroma, 5757–6071)

The eBird team has made the data entry task worthwhile through personally rewarding outputs, and also leveraged the existing community practices around email listservs to provide additional social rewards for data entry. eBird does not displace the well established birding listservs, but created features to work with the existing community infrastructure: eBirders can have checklists emailed to them to forward to friends and birding listservs. Email lists have been the nexus of up-to-the-minute information exchange for years: “even before eBird, one of the major birding things was to provide trip reports...via bulletin boards or mailing lists” (Ceryle, 32622–32850). The ability to forward checklists to listservs also means that contributors are not typing up their observations twice, which further incentivizes use of the system.

A related feature of the site, shared checklists, reduces the data entry burden for groups of individuals who go birding together (Pterodroma, 23939–24333; Ceryle, 34278–34638). Shared checklists allow one person to do data entry for the group while still permitting individuals to adjust their own copies of the checklist. For example, a birding party in Hawaii might collectively record and share a checklist with a dozen Laysan albatross and twice as many Java sparrows, but if one of the observers saw only ten albatross, three dozen Java sparrows, and a Red-tailed tropicbird, that person can adjust her checklist accordingly. This level of control over individual records is important to birders, and shared checklists would be considered less useful if they could not be individually edited to fit personal preferences for record maintenance, speaking again to the intense interest in keeping lists. In addition, an organizer noted that if the birding party includes both eBirders and non-eBirders, the shared checklists have potential to prompt adoption:

There is always one guy in this group sending his sighting to eBird. ... He will share his sightings done during that day to all the 10 people. So they [non-eBirders] will automatically have sightings in their eBird accounts. And this just give them a taste, a taste to follow themselves [continue participating] again and again. I mean, I just opened an account, I already have some sightings, then people just want to follow that. (Diomedea, 1169–12401)

A second feature which is used for social purposes in venues outside of eBird is the ability to share submitted checklists via Twitter and Facebook, added in 2011. While these features do not lead to any social interaction within eBird itself, they make it easier for participants show off eBird checklists in social interactions in other spaces and provide another network-based means for expanding the contributor base. Forwarding emailed checklists to local listservs is part of existing community norms, but as the social media sharing feature is a very recent addition, it is not yet clear what role these technologies may serve in the birding community.

Data entry was also identified as the underlying issue in two other impediments to commitment, specifically among experienced birders: existing birding habits, and concerns about data stewardship. As a project leader noted, the actual commitment is minimal: “once people make the eBird commitment, they’re basically committed to keeping track of birds which they’re sort of doing in their head anyway” (Stercorarius, 48262–48461). Aside from tracking effort data, the additional step of maintaining records and entering data are the main additions to typical birder practices that eBird must convince contributors to undertake. As mentioned earlier with respect to existing habits, many experienced birders already have a system for keeping records. These individuals may be less interested in adopting eBird solely on the basis of eBird’s recordkeeping and listing tools.

Duplication of data entry to turn existing digital records into eBird checklists would pose a substantial stumbling block to commitment from particularly prolific and long-term birders who have amassed a substantial volume of data in other software, so bulk import functionality is available for those who kept their records in Excel or birding-specific software. Some birders, however, also want assurances regarding data stewardship:

Some of the best, very best birders have had a lot of questions about that, how do I know that you vet the records adequately, and why should I bother participating unless you are? Once they can see that, then that kind of is a tipping point. (Stercorarius, 37134–37627)

Presuming these individuals have established recordkeeping systems, contributing to eBird means a commitment above and beyond their current personal data management practices. If assurances of proper data quality management is the factor that convinces these birders to convert into eBirders, their interest in the data stewardship suggests that the scientific merit of the pooled data may be the primary motivation for their contributions. These changes in contribution patterns were specific to expert birders; other changes to birder behaviors are discussed in the following section.

6.7.3 Changing Birder Behavior

The design of eBird, both as a project and as a technology, has clearly had a strong influence on its adoption, which has been observed to lead to changes in the behavior of project participants. The basic expectation from prior research on technology adoption (e.g., Technology Acceptance Model (Davis, 1989), Task-Technology-Fit Model (Goodhue & Thompson, 1995)) is that technology is adopted based on the fit for the people and tasks it supports. As expected, initial adoption of eBird is based on usefulness as a recordkeeping tool. The technology adoption process does not end there, however; once birders recognize the additional personal value that eBird can offer, some change their birding practices to produce more valuable data (Wood et al., 2011; Wiggins, 2011).

In the process of adopting the technology, birders become eBirders, and eBirders do birding differently. Entering observation data online for a citizen science project is an obvious change to previous practices, but the more substantive changes are those occurring in the field. According to a project leader, “eBird wants more than your general birder collects” (Pterodroma, 33454–34240) because the usual recordkeeping practices yield relatively un-specific observations with little information about the effort spent collecting them, which is important for use in scientific analyses.

Improving the data requires following increasingly stringent scientific methods, often a substantial change from recreational birding practices. eBird organizers suggest three changes for better data: submitting complete checklists that include all observed species, contributing counts instead of presence-absence data, and recording effort information about locations, times, and methods. Some birders willingly change the way they bird, recording more information in the field, because eBird provides greater reward for greater effort, creating a “virtuous cycle” in which desirable behaviors are reinforced through a feedback loop. They enjoy increasing benefits with increasing contributions, as their occurrence graphs and

location-specific lists became more accurate and complete: “once they do it they see the value, so they do it more, and so they get more” (Otus, 23059–23223).

Some eBirders are motivated by the reporting and checklist functionality to maintain more regularly collected observations. To make their personal observation data more valuable, a portal organizer reported:

People are keeping track of all the birds they see, they are trying to estimate numbers as best they can, and most importantly, they’re trying to do it on a regular basis. So they are really gearing their birding towards eBird, and eBird rewards them by producing checklists and graphs and maps. (Otus, 15891–16235)

This is a very different aspect of commitment than previously discussed. The prior instances focused primarily on reasons that individuals choose not to use eBird, but for some individuals, once that initial commitment has been made an additional commitment to more intensive participation followed. Techniques to identify these contributors could be valuable for supporting organizing efforts that “focus on bringing out more investment from the people that we have” (Stercorarius, 15078–15182).

Using eBird, and particularly its data visualizations, appears to make the value of using more scientific observation methods self-evident to participants, and birder community practices already reward recordkeeping and data sharing with status and respect. For some birders, these factors lead to a shift in birding behavior that improves scientific outcomes.

The change to behavior is prompted primarily by intrinsically-motivated self-satisfaction and takes several forms. These included satisfaction with the ways eBird supports social recognition or acknowledgement (a traditional aspect of birding culture) through rankings and data transparency, the ability to keep an eye on activity within the birding community by viewing data submitted by others, and access to both personal and aggregate data (Dendroica, Elanoides). The most commonly cited source of satisfaction was the way that eBird enhances the pastime itself, the pleasure that many birders take in keeping lists, and for

some, friendly competition with other birders: “some of the competitive games that people really like, comparing lists with other people’s lists, kinda drives them to engage more, get their lists up-to-date” (Stercorarius, 15183–15428).

eBirders’ satisfaction is further reinforced by the “instant gratification” that eBird provides by immediately updating personal lists and adding personal sightings to public maps and reports within 24 hours of submission. Several organizers reported that being able to access eBird data and reports had changed the way that they and others approach their hobby. A partner project organizer and enthusiastic birder noted, “I birded a different way than I used to bird, because of the way you can enter data into eBird, and then some of it is the gratification of seeing it keep track of things by the county” (Columba, 35210–31007). The particularly powerful intersection of personal interests (bird data) and satisfaction (eBird reports) led to a change in the participation process, and by extension, a refinement of well established community practices that serves both personal and scientific interests. These changes in behavior were also considered to support individual development, discussed next.

6.7.4 Individual Development Through Participation

The interviewees discussed individual development in relation to individuals’ birding skills developing through practice and mentorship (as previously discussed), and the way that using eBird can reinforce good habits, which is the topic of this section. One project leader felt that “on the data entry side, I would say that eBird over time makes general birders much more precise and more aware of what’s happening with their day-to-day birding” (Pterodroma, 36272–36448). This makes intuitive sense given the differences between “conventional” birding and eBirding, which include the development of further attention to scientific detail as the value of the specificity of the data become apparent through system usage (Wiggins, 2011). Individual development through participation, as described by organizers, seems to rest entirely on the development of more scientific birding skills.

Evidence of skill is often considered an appropriate proxy for expertise. A researcher who works with eBird data hopes to quantify contributors' skill levels based on the data they contribute. Any such indicator variables can be incorporated into the scientific data models as a way to control for expected data quality as a function of contributor skill or expertise:

I would like to automatically identify if there are differences in detection rates [of birds] as a function of let's say, your life list, or some species in your life list, and the number of species in your life list for any area relative to everyone else. Or the total number of submissions on eBird, or some sort of involvement, assuming maybe that with more involvement your expertise will go up. (Passerina, 33874–34313)

This quote also highlights the commonly expressed expectation that ongoing participation will lead to further skill growth and individual development.

Surprisingly, while many citizen science projects explicitly hope to educate participants in the scientific method (perhaps prompted in part by available funding sources, such as NSF's Informal Science Education programs), relatively few seem to place substantial emphasis on domain-specific learning and skill development as a desirable outcome for participants. This is an interesting incongruity, as there is frequent concern over participant skill and expertise as relates to data quality, but domain learning seems to be considered only a means to an end. Project evaluators for the Citizen Science Online grant also noted that enabling participants to conduct their own inquiries using the eBird data is generally “seen more as a matter of personal enrichment than explicitly providing a platform for amateurs to produce professional research results or engage in advocacy (while not ruling out either of those uses)” (Allison-Bunnell & Thompson, 2007, p. 5). From a participant's perspective, however, developing further domain expertise may be far more motivating than learning about the scientific method or producing professional research.

Interviewees noted that the ability to visualize the large volumes of data that eBird has accumulated using animated range maps, produced with high-performance computing

resources from TeraGrid, can lead to epiphanies regarding the relationships between species and habitat. They speculated that the visualizations made the scale and value of eBird data more accessible to casual audiences. Showing non-scientists the connections between habitat preservation and bird distribution, for example, was highlighted as a potentially transformative experience resulting from access to sophisticated data visualizations.

They mentioned that the animated migration maps, in particular, were excellent at drawing attention. A Lab staff member discussed a presentation by an eBird project leader at a bird festival, saying:

The biggest news in the eyes of the audience, that impressed them the most, is that we're able to begin to show data...so you could see migrating species. ... And for them to see that data was really exciting to a lot of people, and to understand that you can study that data. Are the birds coming earlier because of global warming? Are they leaving earlier, staying longer? You know, what's really going on? (Elanoides, 12942–14021)

The quote also suggests that these visualizations prompt the development of hypotheses and research questions among non-scientists by helping them see the potential uses of the data.

Although it is not a project output that is explicitly educational, being able to access these data was perceived to be a meaningful and potentially transformative experience. A project organizer connected the ability to make large scale, complex data accessible through visualization to the Lab's mission:

The fact is that to see the dynamics, spatially and temporally, of how these things change, and to do it at such a broad scale as we can do it, is transformative in the way people think about biodiversity and natural history. All that kind of visualization of data is very important, and serves the Lab's mission. (Dendroica, 27494–27823)

For many participants, the data visualizations may be just another way to find the information they desire, but for others, seeing the aggregated data in a different way can stimulate a change in the way they understand the relationship between biodiversity and habitat preser-

vation.

6.8 Answering the Research Questions

Returning to the research questions, the eBird case study provides insight into the primary constructs of virtuality, technology, participation, organizing, and scientific outcomes, each of which is summarized in this section.

Virtuality is a fundamental characteristic of project organizing and participation. It benefits eBird by allowing greater participation on a global scale through online data submission, which has proven critical to supporting increasing scale of participation. The distributed project management structure also supports outreach efforts because project leaders can more easily attend in-person events that are clearly important for participant recruitment across a wider geographic scale, which further increases virtuality. Virtual contribution is supported by a sophisticated technological system which has a substantial impact on participation and organizing.

Technologies are a core input and product for eBird that supports the organizing of virtual participation. The impacts of technology on participation and organizing were most apparent in the complex interactions of technology design with community practices and science processes (and by extension, scientific outcomes). The discussions of institutional relationships revealed that technology adoption of eBird by third parties as infrastructure signals increasing trust and dependency on the system, but also highlighted potential issues related to data quality due to adaptation (or appropriation) to fit existing protocols into eBird's data management structures. The interviewees further stressed the importance of swift system feedback in encouraging ongoing participation, as well as the impact of visualization on lay people's ability to grasp the larger picture and begin to understand the importance of habitat for bird conservation. In addition to sustainability of human resources

in the form of contributors, eBird's sustainability is intimately connected to the technology through provision of services such as portals and kiosks, and the use of grant funding for innovation rather than operating costs. As project participation grows, however, the reliance on revenue streams tied to the system itself could lead to ongoing challenges with scalability that may require additional investment in technology.

The technologies that support eBird's participation processes demonstrate an alignment of scientific and personal interests which has produced a system that incentivizes participation by producing outputs that satisfy contributors. Basing the participation protocols on minor changes and extensions to existing community practices nearly eliminates any formal training requirements related to the core task of bird observation (although presentations provide an informal type of training in use of the system) and makes it easy for birders to integrate into their established routines. The inclusion of features that reduce data entry burden further supported social participation in the birding community while passively leveraging contributors' personal networks and community connections to encourage increasing adoption. The participation processes permit participants to contribute at any level with which they are comfortable by supporting multiple protocols, which is relatively unusual in citizen science more broadly but appears to promote participation by a wider audience.

Organizing efforts used gentle encouragement to adopt more rigorous protocols, combined with the affordances of the technology itself, which has led to changes in birder behavior that generate higher quality scientific data. An important enabler of these shifts in long-standing habits is that by design, following more scientific protocols also benefits the birders who enjoy recording and exploring bird data. eBird's organizers have also maximized the value of participant contributions through role expansion, recruiting a network of the most expert and committed contributors to help with data quality management as meta-contributors. These observations called attention to the importance of communication to promote partic-

ipation and encourage improved data quality, but more interestingly, showed that indirect communication through system feedback is a valuable complement to direct interaction between project leaders and participants. Both communication and participation led directly to individual development, primarily with respect to domain skills, as did access to data and visualizations that were reported to have transformative potential.

Successfully organizing large-scale *participation* was largely attributed to the strategic choice to hire project leaders drawn from the birding community, which has had a substantial influence on both technology and participation. Notably, the project leaders are not only birders, but are known and respected within community, so their reputations and networks directly benefit eBird. While they have science backgrounds, the project leaders were birders first, and their close connection to the birding community has impacted many aspects of the project, an influence most apparent in the development of the birder-centric system features that are eBird's primary attraction for participants. The project leaders took advantage of the existing social network of birders with community-based recruitment strategies, primarily through direct outreach to indigenous organizations in their organizational field, which further enabled the small staff to produce large-scale results. This approach was also employed for organic growth as eBird extended its global contributor base, relying on local organizers around the world to arrange reviewer networks and recruit participants. Engagement in both global and local partnerships pointed to the influences of eBird's institutional context and partners on both technology and outputs, as partnerships bring additional resources to the project but require adequate organizational and staff capacity to organize.

Scientific outcomes have been one of eBird's strengths, as the project has demonstrated that birders can contribute scientifically useful data, which have been used in scholarly publications. The project's effective organization and high levels of participation have yielded a large volume of data. One of the primary reasons that this was possible is the project's policy

to accept all data that participants wish to contribute (and supporting these contributions with appropriate tools), and expecting researchers to select a subset of the data that suits their scientific interests, rather than expecting all participants to follow a rigorous, detailed scientific protocol that would have suppressed participation. Both the volume of data and its particular characteristics (such as geographical bias) were identified as meaningful influences that forced researchers to take a different approach to data analysis that involved developing new statistical models and utilizing high-performance computing resources. Despite (and because of) these adaptations to science processes, the data have had far-reaching benefits for science-based decision support and cross-disciplinary research applications.

eBird's initial and ongoing user-centered technology design processes are combined with community-centered participation processes, but this foundation was not adequate to ensure success. The critical factors that proved to be the tipping point for the project were: 1) changes to the way that the project was organized, particularly the addition of project leaders who are respected and connected in the birding community, and 2) the shift in the technology design to better satisfy the personal interests of the community through improved feedback to contributors. The discussion of the eBird case demonstrates that each of the constructs in the research questions are tightly interwoven in this large-scale technology-driven citizen science project.

CHAPTER VII

Theoretical Framework

This chapter discusses the theoretical framework that was iteratively developed throughout the study. It served as a lens for focusing the research and, in modified form, also became an output of the research process. It identifies several practical considerations for citizen science projects and can help direct future research. This chapter presents a systematic review of each concept, with examples drawn from the cases.

7.1 Theoretical Framework

The theoretical framework presented in this chapter (Figure 7.1) is one of the contributions of the current work. It represents an iteratively-developed, empirically-grounded framework of citizen science that is both congruent with prior models and more refined than existing frameworks. The diagram shows the expected relationships between inputs, moderators, and outputs of a citizen science project. It also shows the concepts within each of these categories that were identified as having a meaningful impact on project organizing, participation, and outcomes, corresponding to both the research questions and emergent themes from the data.

In Figure 7.1, rectangles represent project-level concepts and ovals represent individual-level concepts. In addition, while connections are not shown between the individual concepts within in each category, many of them are related to one another. For example, the states of satisfaction and commitment are connected both logically and in the interview data.

Evidence of each of these concepts appeared in the data, and many were further supported by the complementary conceptual models that will be introduced in this chapter. Table 7.1 summarizes the concepts from the theoretical framework. Notably, *all* concepts are based on empirical evidence in addition to any relevant conceptual models, and several concepts attributed to complementary conceptual models were already present in the initial framework, as discussed in the following section.

The initial framework discussed in Chapter II served as a set of sensitizing concepts for data collection and analysis. It was based on an inputs-moderators-outputs-inputs (IMOI) structure. Through several cycles of revision to incorporate insights gained through empirical observations, the current framework evolved through expansion, simplification, and reconfiguration of earlier versions (e.g., Wiggins & Crowston, 2010). The deductive process of theory development was informed by inductive coding and analysis.

Concept	Category	Level	Brief Description	Origin
Environment	—	Global	Broader social, cultural, economic, & infrastructural conditions	Ganz (2000)
Scientific interests	Inputs	Project	Research focus, questions, & goals	Shirk et al. (2012)
Community	Inputs	Project	Social group with shared interests & practices	Initial framework
Resources	Inputs	Project	Financial, organizational, material, & human assets	Initial framework
Institutions	Inputs	Project	Institutional context & field	Data
Mission	Inputs	Project	Project goals & intentions	Ganz (2000)
Technologies	Inputs	Project	Tools & processes for reaching goals	Initial framework
Biography	Inputs	Individual	Life experience	Ganz (2000)
Personal interests	Inputs	Individual	Motivation, goals, & values	Shirk et al. (2012)
Skills	Inputs	Individual	Knowledge, expertise, & abilities	Ganz (2000)
Networks	Inputs	Individual	Personal contacts	Ganz (2000)
Science	Processes	Project	Research procedures	Initial framework
Design	Processes	Project	Initial & ongoing project decisions	Data
Organizing	Processes	Project	Project & volunteer management	Initial framework
Participation	Processes	Individual	Taking part in project activities	Initial framework
Sustainability	States	Project	Ability to continue pursuing goals	Initial framework
Commitment	States	Individual	Ongoing responsibility or obligation	Initial framework
Satisfaction	States	Individual	Fulfillment	Initial framework
Contributions	Outputs	Individual	Task & non-task products	Initial framework
Individual development	Outputs	Individual	Learning & socialization	Shirk et al. (2012)
Scientific knowledge	Outputs	Project	Scientific findings	Initial framework
Broader impacts	Outputs	Project	Intended & unintended products beyond scientific findings	Shirk et al. (2012)

Table 7.1: Summarized concepts from the theoretical framework.

Throughout the development of the framework, congruence with existing theories was

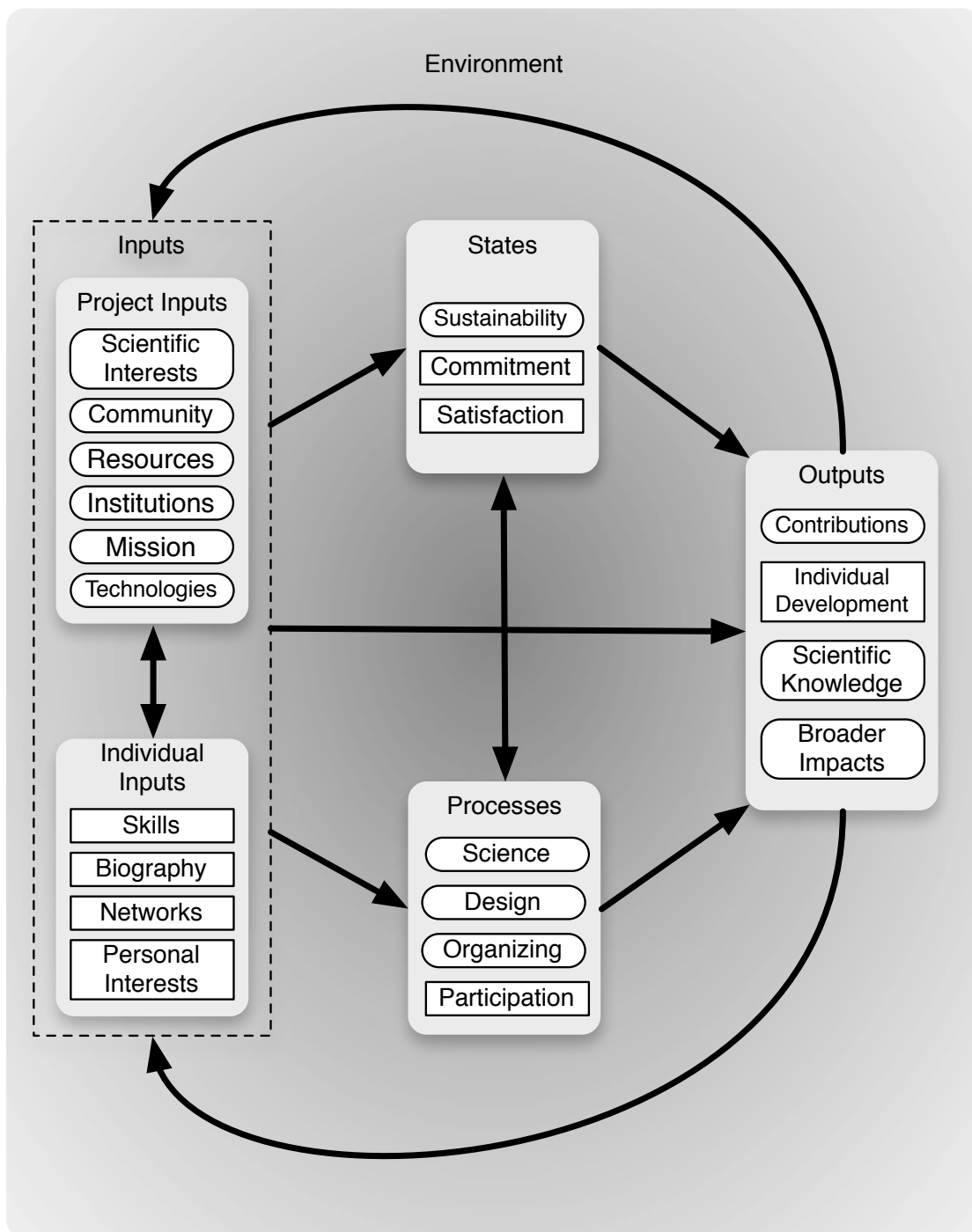


Figure 7.1: A theoretical framework of citizen science as a type of virtual organization.

repeatedly examined, leading to the inclusion of elements from two complementary theories drawn from other disciplines, discussed next. These concepts captured the essence of the empirical observations related to several facets of the case studies, integrating elements of models that represents citizen science from a program evaluation standpoint and from the perspective of collective action. The description of the framework then turns to the inputs, processes, states, and products, covering each concept in turn and drawing examples from the case studies.

7.2 Complementary Theories

Throughout data collection and analysis, as the theoretical framework was tested, elaborated, and refined, additional theories from the literature were evaluated for complementarity to the developing theory. Two theoretical models were particularly helpful, as they validated several empirical observations and provided useful ways to frame the concepts that emerged from the data. The Deliberate Design Model for citizen science is drawn from the ecology literature, and an early version of the theoretical framework for this study contributed to its development. As one of very few other conceptual models representing citizen science, its congruent structure and complementary concepts enhanced the evolving theoretical framework in this study.

The Strategic Process Model comes from the sociological literature focused on social movements theory and collective action, which is particularly apropos to this phenomenon because from a theoretical standpoint, organizing and participation processes in citizen science can be represented as a form of collective action. The Strategic Process Model provided a conceptualization of individual leadership qualities that forms the basis of the individual-level inputs. The process orientation and relevance to organizing large numbers of voluntary contributors made the Strategic Process Model helpful for representing aspects of citizen

science that were evident in the data but had not been represented in other conceptual frameworks.

The Deliberate Design and Strategic Process models are briefly introduced here and their contributions to the theoretical framework are described.

7.2.1 Deliberate Design Model from Ecology

The Deliberate Design Model from Shirk et al. (2012) (see Figure 7.2) is itself an adaptation of a general model, much like the IMOI model. Developed in the context of ecology research, it represents a particular perspective of citizen science, and is notable for being the only other theoretical framework besides that presented here that aims to describe citizen science as a phenomenon. The Deliberate Design Model draws on the program logic model from the W.K. Kellogg Foundation (also known as *program theory* in the evaluation field), which was created as a framework for program development and evaluation (Kellogg:2004, 2004).

Structurally, the program logic model is a set of sequential steps in a linear model: Resources are established prior to Activities, Activities produce Outputs that yield Outcomes, and Outcomes eventually produce Impacts. These are roughly equivalent to the IMOI model categories for Inputs, Moderators, and Outputs, but discriminates between first- and second-order products of the project as well as long-term organizational, community, and/or system level changes. The basic model as developed in Kellogg:2004 (2004) is entirely linear, but the Deliberate Design Model acknowledges the cyclic properties of feedback with outputs serving as inputs to ongoing projects, demonstrating further congruence with the IMOI model structure.

The correspondence between concepts in the Deliberate Design Model and the theoretical framework developed in this study is shown in Table 7.2. The specific rationale for these correspondences are discussed in more detail later in this chapter. The inputs for the Deliberate

Original Concept	Corresponding Concepts
Inputs: Scientific interests	Inputs: Scientific interests
Inputs: Public interests	Inputs: Personal interests
Inputs: Identify question or issue	Processes: Science & Design
Activities: Develop infrastructure	Processes: Design
Activities: Manage project implementation	Processes: Organizing
Outputs: Observations	Outputs: Contributions
Outputs: Experiences	Outputs: Individual development
Outcomes: Science	Outputs: Scientific knowledge
Outcomes: Social-ecological systems	Outputs: Broader impacts
Outcomes: Individuals	Outputs: Individual development
Impacts: Conservation	Outputs: Broader impacts

Table 7.2: Correspondence of concepts from the Deliberate Design Model.

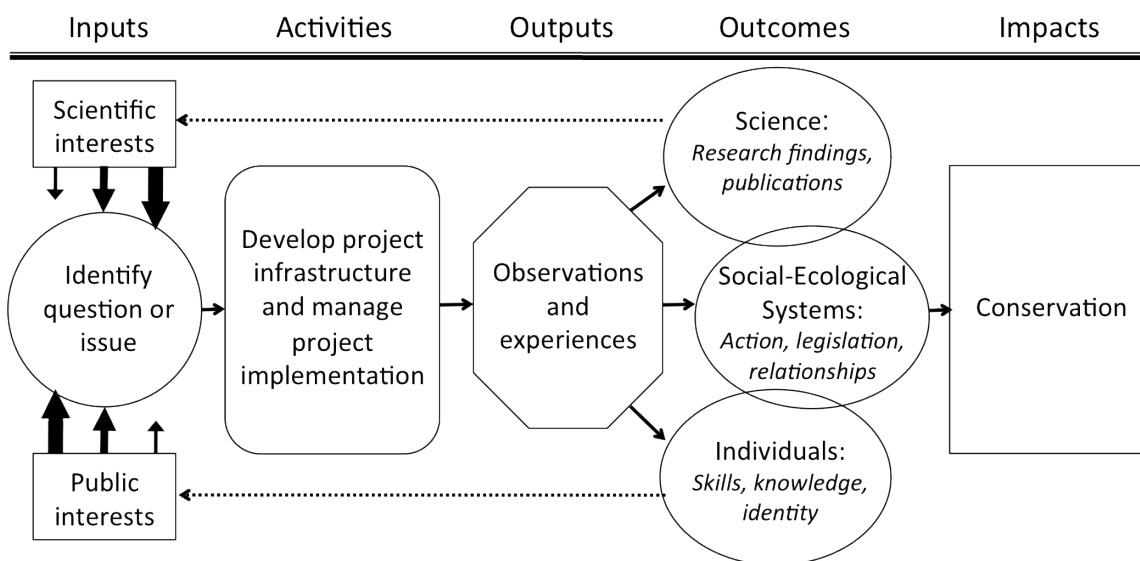


Figure 7.2: The Deliberate Design Model from Shirk et al.

Design Model are minimal, focusing primarily on the interests that motivate the creation of a project, and is intended to highlight the questions of “whose interests are being served?” The Deliberate Design Model confirmed the empirical observations of scientific and personal interests from this study, and provided a useful way to describe the goals and motivations of different groups for engaging in citizen science. These two concepts proved particularly beneficial for identifying the importance of the alignment of scientific and personal interests in each case.

The Deliberate Design Model also highlights the different types of outputs observed in the data, focusing again on whose interests are served, into categories that are reflective of different potential audiences or purposes for a citizen science project. Science outcomes are a category of outputs that were retained intact as scientific knowledge in the framework, as it was an existing point of agreement with the prior version of the framework. The concept of broader impacts encompasses the potential outcomes similar to those represented under social-ecological systems, but from a broader perspective.

7.2.2 Strategic Process Model from Social Movements Theory

As mentioned above, citizen science can be viewed as a form of collective action. However, collective action in citizen science departs from the political focus to which this term is typically applied. Social movements theory focuses primarily on mobilizing structures, political opportunities, and framing processes in order to understand emergent collective action (McAdam & Scott, 2005). In this literature, collective action is generally conceptualized as a response to oppression, but there is no such parallel in citizen science. Instead, participation by members of the public in doing scientific work represents cooperation and collaboration.

The notion of opportunity structures provides a means for bridging contexts of contention and cooperation. In social movements theory, political opportunity structures enable the acquisition of resources, and organizing processes generate greater value when mobilizing

resources from a wider range of sources (Ganz, 2000). A complementary view of mobilizing structures, defined as “collective vehicles, both formal and informal, through which people come together and engage in collective action” (Adam, Tarrow, & Tilly, 1997, p.155), links mobilizing structures with opportunity structures.

Translated to the context of citizen science, these mobilizing structures can be understood as facilitative opportunity structures through which conditions conducive to public participation are created and resources are mobilized. Several inputs and processes (e.g., institutions, technology, organizing) combine to generate facilitative opportunity structures that support cooperative collective action. For example, ICT enable broader participation by facilitating input from a wider range of contributors, enhancing the resources available to a citizen science project. This view harmonizes political opportunity theory focused on contexts of contention with other forms of collective action—such as citizen science—that instead arise from cooperation.

In the language of the collective action theorists, facilitative opportunity structures are created (in part) when rules for participation are created by project organizers in the form of protocols, which provide the rules for individual observations. These protocols act as “collective choice” actions made by organizers that reflect a type of governance structures for participation; contributors who carry out the protocols engage in “operational choice” actions (Ostrom, 1990). Operational choices lead to actions (observation) and outputs (data) that are subsequently united into a collective output. Citizen science organizers are therefore creating facilitative opportunity structures that rely on pooled interdependence to mobilize distributed actors. These structures of organizing and participation create a different form of collective action from those characterized in the literature on organizations and movements.

While several frameworks and models from social movements theory may provide new insights into citizen science, the Strategic Process Model includes several concepts that de-

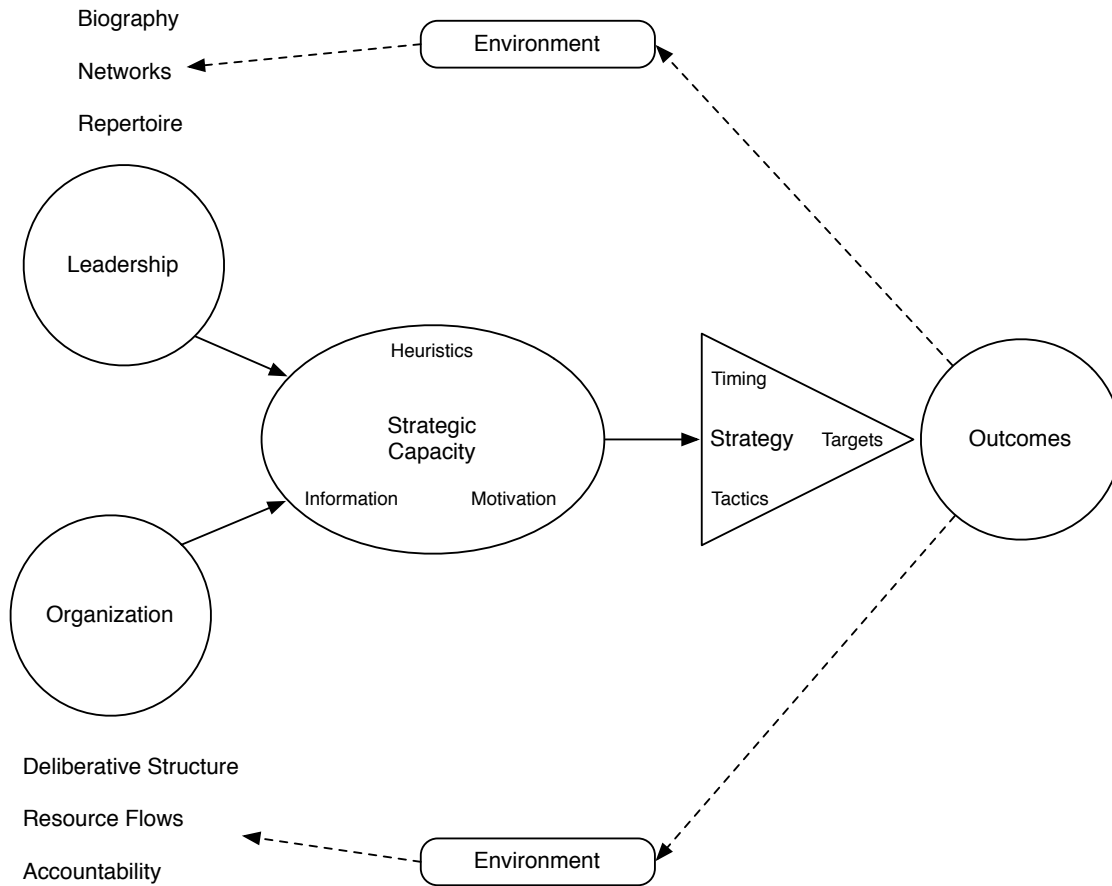


Figure 7.3: Reproduction of the Strategic Process Model from Ganz (2000).

scribed the emergent themes in the empirical data. The Strategic Process Model (see Figure 7.3) was developed to conceptualize the notion of strategic capacity in collective action (Ganz, 2000). Ganz defined strategic capacity as the likelihood that an organization will develop effective strategy, and his model identifies links between leadership and organizational variables that create conditions conducive to developing strategic capacity.

Like the IMOI model and the program logic model, the Strategic Process Model shows a link between inputs and outputs, but is substantially different in focus, building on social psychology, cognitive psychology, and organization theory. The correspondence of concepts between the Strategic Process Model and the theoretical framework for citizen science is shown in Table 7.3, and the rationale behind these relationships will later be discussed in

Original Concept	Corresponding Concepts
Environment	Environment
Leadership: Biography	Inputs: Biography
Leadership: Networks	Inputs: Networks
Leadership: Repertoire	Inputs: Skills
Organization: Deliberative structure	Inputs: Institutions
Organization: Resource flows	Inputs: Resources
Organization: Accountability	Inputs: Institutions
Strategic capacity: Heuristics	Inputs: Resources
Strategic capacity: Information	Inputs: Resources
Strategic capacity: Motivation	Inputs: Mission
Strategy: Timing	Processes (all)
Strategy: Targets	Processes (all)
Strategy: Tactics	Processes (all)
Outcomes	Outputs (all)

Table 7.3: Correspondence of concepts from the Strategic Process Model.

further detail. The key elements of the Strategic Process Model that were adopted for the theoretical framework in this study include the attributes of environment and qualities of leadership. The specific concepts related to organization were already present in the theoretical framework as institutions and resources. The elements which were not explicitly included were strategic capacity and strategy. Two of the concepts making up strategic capacity were already encompassed by the project-level input of resources; mission was included as a translation of motivation at the aggregate level. The components of strategy, translated to the context of scientific work rather than opposition to injustice, were related to aspects of the project processes.

Ganz identifies biography, networks, and repertoires as crucial variables related to leadership. The empirical evidence from the case studies showed that it was not leadership roles that were important in citizen science so much as the individual qualities associated with leadership. Leadership is generally associated with specific roles, but the qualities of leadership are present, to varying degrees, in every person regardless of role. The aspects of leadership identified by Ganz provides an elegant representation of a range of concepts existing in a prior version of the theoretical framework. They also allowed the conceptual

separation of these personal characteristics from processes related to leadership which are considered a part of organizing processes.

These three variables (biography, networks, and repertoires), associated with leadership in collective action contexts, are therefore applicable to the full breadth of individuals participating in citizen science, and the concepts of biography and networks were adopted without modification. Repertoires, however, represents the knowledge of skills and behaviors applicable to collective action; elsewhere, they are also conceptualized as cultural knowledge (Williams, 1995), which is particularly relevant to communities of practice. In the theoretical framework for citizen science, these concepts are more simply labeled “skills,” as Ganz associates repertoires with competence and local knowledge. Competence and local knowledge can be assets for both citizen science and other related phenomena by providing a foundation of contextual knowledge that supports participation and improves outcomes.

Additional contributions to the theoretical framework drawn from Ganz’s model include the concepts of mission and environment. Ganz represents motivation as a variable related to strategic capacity. This concept is labeled as mission and included at the project level in the theoretical framework for reasons discussed in the following section.

Ganz devotes little attention to the concept of environment, but notes that it represents a broader set of conditions (typically social and political) external to the organization which are ever-changing, set the stage for collective action, and are reshaped by actors’ strategies. A similar concept is also present in the collective action models from McAdam (1999), where it is labeled “broader socioeconomic processes” and framed in the context of political insurgency. This notion refers to the “structure of political opportunities” that accumulate over time, leading to expanded opportunities for collective action. As noted above, this is a parallel concept to that of facilitative opportunity structures in the context of citizen science.

In the current theoretical framework, environment represents broader social, cultural,

economic, and infrastructural conditions that provide opportunity for collective action. For example, an example of a characteristic of the environment is public sentiment. While positive or negative public sentiment towards scientific research undoubtedly influences citizen science projects, these effects are indirect and not evident in the data collected for this study. Environment is therefore included as the backdrop to the concepts in the framework, as they are all affected by these conditions but influence them only indirectly.

In the following sections, the concepts from the theoretical framework are discussed in more detail. Some of these concepts were introduced earlier in Chapter II, and are now presented with additional discussion of the rationale for their inclusion in the framework.

7.3 Project Inputs

Inputs such as scientific interests and communities (among others) are the basic building blocks for a citizen science project. Each project requires different proportions of these inputs but needs similar assets (Table 7.4). The inputs discussed in this section represent the resources and conditions of a project at the aggregate level. These inputs typically serve as assets for organizers, and less directly for contributors, but can also create constraints. Balancing these elements is a primary consideration in initial project design, and also influences ongoing operation.

Concept	Brief Description	Origin
Scientific interests	Research focus, questions, & goals	Shirk et al. (2012)
Community	Social group with shared interests & practices	Initial framework
Resources	Financial, organizational, material, & human assets	Initial framework
Institutions	Institutional context & field	Data
Mission	Project goals & intentions	Ganz (2000)
Technologies	Tools & processes for reaching goals	Initial framework

Table 7.4: Summary of project inputs from the theoretical framework.

7.3.1 Scientific Interests

Scientific interests are a necessary ingredient for a citizen science project; projects without scientific interests, questions, or goals are (perhaps arguably) simply a form of outreach, education, or scientific communication. While that style of interaction can certainly have value, the focus in this study was on projects that intended to produce scientific knowledge. Scientific interests can include hypothesis-driven research, ongoing monitoring, and decision support where science-based intervention or policy is needed.

The concept of scientific interests was drawn from the Deliberate Design Model (Shirk et al., 2012), which highlights the importance of alignment between the interests of organizers and contributors. While seemingly self-evident, this alignment is not a given, though it has been identified as a logical way to help ensure sustainability of project participation (Wood et al., 2011). Researchers and project organizers must be sensitive to the interests of contributors and the ways in which both scientific and personal interests are mutually served by participation activities. For the GSP, the addition of new plants for monitoring in response to requests from contributors enabled LeBuhn to consider additional research questions, demonstrating an alignment of scientific interests and personal interests.

7.3.2 Community

As defined by Wenger (1999), a community of practice is a social group of individuals who interact with and learn from one another, engage in joint activities based on their shared interests, and have shared repertoires of experiences, knowledge, and skills. While communities were generally external to citizen science projects in this study (although communities can certainly form around projects) they can play an important role in supporting participation and organizing. Communities do not necessarily occur at the project level, but existing communities are often leveraged by projects, leading to their categorization as

project inputs.

In the cases presented in this study and more generally, citizen science projects tend to seek contributors from an existing community of practice whose interests are aligned with the scientific interests of the project. Often such a community will have associated institutions and established practices which can serve as resources as well. For example, eBird focuses on birders, Great Sunflower Project appeals to gardeners, and MountainWatch works with hikers. For each of the cases, these primary communities of practice are hobbyist, enthusiast, and/or leisure communities in which individuals become involved out of personal interests.

7.3.3 Resources

Resources are often in short supply for citizen science projects, which are fundamentally reliant on human resources to achieve their goals. Resources are not necessarily up-front pools of assets upon which organizers can draw, but may require organizing processes to identify, appropriate, and assemble (Rao, 1998). One of the reasons for forming citizen science projects within the boundaries of an organization (discussed below) is that existing organizations can provide resources needed during the initial stages of project formation that are otherwise inaccessible. As previously discussed, access to resources in citizen science is in part an outcome of the development of facilitative opportunity structures that arise from the interaction of such elements as environment, institutions, and organizing.

From a more operational perspective, resources refer to financial, organizational, material, and human assets that are leveraged or required for project creation and operation. Even projects with strong funding found that resource constraints were a perennial issue. At the opposite end of the spectrum of available resources, citizen science holds the promise of accomplishing greater outcomes for the fiscal and organizational resources invested in them. This expectation is primarily because citizen science relies heavily on “free” human resources in the form of volunteer contributors.

The human resource dependency of citizen science projects is a definitional characteristic. If a relatively large number of contributors (compared to available professional human resources) were not required to address scientific interests, the citizen science approach would be an unlikely choice. The degree to which successfully achieving project goals depends on participation is often a matter of scale, where scale typically refers to spatial or temporal ranges, or volume of data. Economies of scale that are often attributable to technologies make access to volunteers as human resources more feasible than ever before. As a result, the decreasing costs of technologies is a powerful enabler of the rapid spread of citizen science.

Nonetheless, as the cases showed, limits on human resources in the form of project organizers was the primary constraint on project growth. For example, Mountain Watch organizers encountered challenges with a “lack of resources to put time into setting up plots and making sure someone is checking them” (Geum 2011, 13054–13250).

7.3.4 Institutions

The contexts of most citizen science projects involve a variety of institutional and organizational influences, two aspects of which are particularly relevant to these cases: the institutional environment and the organizational field. Organizational theorists from the “new institutionalism” paradigm variously define institutions, often focusing on such themes as competition, conflict, and change (Powell & DiMaggio, 1991). Across disciplines such as economics and political science, institutions have been conceptualized as frameworks of rules, processes, and social structures (Shepsle, 1989); social arrangements and governance structures that are intended to reduce transaction costs (Williamson, 1981); and assemblages of rules that prescribe actions (Ostrom, 1986). These conceptualizations of institutions take a variety of perspectives, seeing institutions as based upon equilibria, norms, and rules, depending on different assumptions (Crawford & Ostrom, 1995).

The rule-based view of institutions is most useful for understanding citizen science, as it

does not rely on assumptions of rationality or shared perceptions of behavioral norms, both of which are difficult to apply to distributed collaboration carried out through voluntary participation. The definition of an institution from a rule-based perspective as “human-constructed constraints or opportunities within which individual choices take place and which shape the consequences of their choices” (McGinniss, 2011, p.170) highlights the interaction of social structures and individual choice. This view is pertinent to citizen science, in which socially-constructed scientific research expectations intersect with organizational practices, as it recognizes these influences on the decision-making processes of design, organizing, and participation.

Institutional Environments and Organizations

Institutions make up the environments in which organizations operate, and organizations therefore draw upon their institutional environments as sources of knowledge, resources, and supporting structures (Scott, 1991). Organizations are likewise variously defined, although most definitions agree with the basic conceptualization of organizations as “social structures created by individuals to support the collaborative pursuit of specified goals” (Scott & Davis, 2007, p. 11). These social structures are reproduced by structuration processes in which ongoing patterns of action based on the rules and resources provided by the institutional environment (Giddens, 1979).

This general definition of organizations as structure and process is most useful to the discussion of citizen science, as the institutional environments and organizational forms in which they operate vary substantially. The typical institutional environments in which citizen science projects are organized include nonprofit organizations, academic institutions, and government agencies, all of which are subject to different institutional environments. In a few unusual instances, citizen science projects have also been organized by corporate entities, such as Microsoft’s Pathfinder (Luther et al., 2009) or IBM’s involvement in Creek-

watch (Kim et al., 2011). The cases selected for this study include projects founded within conservation-focused nonprofit organizations (eBird and Mountain Watch) and a university (the GSP). While some citizen science projects are founded outside of a formal organization, this appears to be relatively rare, likely due to the advantages conferred by institutional environments, organizational resources, and interorganizational relationships.

Organizational Fields

Interorganizational relationships shape the broader organizational field, defined as “a collection of diverse types of organizations engaged in competitive and cooperative relations,” (Scott & Davis, 2007, p.117). In the cases presented in this study, the organizational field includes organizational and research partners, funders, and other institutions that interact with the project, as well as other citizen science projects. Although it is infrequently acknowledged by practitioners and related literature, citizen science projects do compete for funding resources and participants, particularly in cases where the domain focus and protocol are similar enough that differentiating between the projects can be challenging.

At the same time, some citizen science projects collaborate with others, co-promoting participation in similarly-themed projects that have different levels of complexity in their participation protocols, are active in different geographic regions, or require different levels of skills and types of knowledge. For example, one respondent to the survey discussed in Chapter III mentioned eBird as a competitor to their own bird monitoring project, which had preceded eBird but was subsequently outperformed by it on several dimensions. By contrast, another respondent to the same survey works with eBird as a collaborator in a mutually beneficial arrangement. However, as prior research has noted, compensating for shortages in internal competencies is not the sole reason behind interorganizational collaborations because these ongoing relationships also strengthen internal skills through network-based learning (Powell, Koput, & Smith-Doerr, 1996).

Formal partnerships between and among organizers and institutions are quite common in citizen science projects. Every case in this study had multiple partnerships with other citizen science projects, individual collaborators, nonprofit organizations, research teams, federal agencies, schools, and/or other groups. The capacity of a project to form, leverage, and benefit from such arrangements appears to be a useful indicator of project health, and could potentially help in predicting future performance. Interviewees often noted the additional effort required to organize and maintain such institutional relationships, but also the benefits of shared resources or access to needed expertise and skills. The formality of partnership arrangements in the cases ranged from contractual obligations and funding relationships to handshake agreements and informal in-kind provision of staff time and services.

In all three cases, the organizational field evolved over the life of the project, beginning with the initial organizational affiliations of the organizers (e.g., Cornell Lab of Ornithology, Appalachian Mountain Club, San Francisco State University) and expanding to include a growing array of partners that brought both resources and learning opportunities. Loose partnerships became formalized relationships in some circumstances, but all of the projects maintained a mix of both formal and informal interorganizational relationships. As mentioned in the eBird case study, these partnerships brought both opportunities and constraints, but the project leaders continued to maintain a close match between formal partnerships and organizational mission, discussed next.

7.3.5 Mission

Ganz (2000) uses the term motivation to describe an aspect of strategic capacity in his Strategic Process Model, but the theoretical framework for this study represents the underlying concepts with the term mission for two reasons. First, motivation is an individual-level concept, incorporated in the theoretical framework for citizen science under personal interests, while mission is a parallel aggregate-level concept often employed in organizational

contexts, but nonetheless applicable to projects and generally represented as goals. Second, interviewees repeatedly referred to organizational mission in the two cases that are set in nonprofit organizations, and although the concept of a mission is more of an institutional-level conceptualization, it is a broader label that can also encompass project goals. For example, a Mountain Watch organizer related project goals to organizational mission by saying, “we are not trying to do citizen science for citizen science. We’re trying to do something that is helpful to our science department and our mission” (Geum 2011, 7072–7337).

From a conceptual standpoint, mission need not be clearly articulated or aspirational to represent project-level motivations. However, organizers in all three cases specifically referred to mission—organizational, project-specific, or both—as a guide for decision-making. Project organizers reported using organizational mission along with project goals in the design and organizing processes. For projects conceived within a nonprofit organization, leveraging an established mission is only natural, and the project goals for the cases in this study were crafted in alignment with organizational mission. The project mission and goals were frequently mentioned in conjunction with broader impacts, highlighting ways that project outputs were transformed into unexpected but mission-relevant outcomes. Mission was also connected to participation, satisfaction, personal interests, scientific interests. In the eBird project, for instance, project organizers linked large-scale data analysis and visualization of contributor data to organizational mission: “to see the dynamics...is transformative in the way people think about biodiversity and natural history. All that kind of visualization of data is very important, and serves the Lab’s mission” (Dendroica, 27461–27823).

7.3.6 Technologies

Although the definition of technology has historically related to knowledge and practice, as discussed in Chapter I, in more modern usage, technology refers to the tools and processes

used to accomplish a goal. In the theoretical framework, technologies are broadly conceived from a social science perspective. For example, the term can pertain to information and communication technologies such as computers and GPS devices, but the concept can also encompass the paper data sheets used in the field. Both of these meanings have been employed in the foregoing discussions of technologies used to support citizen science.

Information technologies are substantially more important in technology-supported citizen science projects like eBird and GSP than the preceding models of volunteer monitoring that had little reliance on ICT, of which Mountain Watch is more representative. Notably, ICT can also be a fundamental enabler of citizen science, as seen in eBird and GSP, as well as a mediator of participant and organizer interactions. End-user information technologies were rarely mentioned by interviewees, as it was an implicit assumption (and in fact a prerequisite) that contributors have access to the minimum required technologies of a computer and the communication infrastructure provided by Internet access. As citizen science becomes increasingly technology-dependent, it relies not only on project-level technologies but also the infrastructure and personal computing technologies available to participants.

Across the cases in this study, technologies were most strongly linked to design processes, although resources were also a commonly related theme. Technology design and use is of particular interest given the potential of cyberinfrastructure to support citizen science (Chin & Lansing, 2004). Early examples of systems intended to provide cyberinfrastructure for citizen science include a platform for managing invasive species monitoring projects (Graham et al., 2008) and the National Geographic FieldScope project to develop a collaborative geospatial platform for citizen science (Russell, Switzer, & Edelson, 2011), but information technologies specifically adapted to citizen science are scarce. Best practices guides recommend that project partnerships include a scientist and an educator to address the scientific and educational goals of the project, and a technologist to address potentially substantial

data management and information systems challenges (Bonney & LaBranche, 2004). Understanding the range of interactions between diverse end users and technologies that support the scientific research is important to creating usable, robust systems for collecting useful independent contributions by distributed volunteers (Luther et al., 2009).

7.4 Individual Inputs

Each individual brings personal background, characteristics, and resources to a citizen science project (Table 7.5). Staff and volunteers have diverse demographics, levels of skill, and motivations for participation that influence their individual contributions to the project. The types of individual-level inputs discussed in this section are not unique to either project organizers or participants, but apply to all contributors and users of data generated by a citizen science project. Three categories of individual inputs are related to leadership qualities that everyone possesses in varying proportions, while personal interests encompasses several inter-related concepts such as motivation and personal values.

Concept	Brief Description	Origin
Biography	Life experience	Ganz (2000)
Personal interests	Motivation, goals, & values	Shirk et al. (2012)
Skills	Knowledge, expertise, & abilities	Ganz (2000)
Networks	Personal contacts	Ganz (2000)

Table 7.5: Summary of individual inputs from the theoretical framework.

7.4.1 Biography

The theoretical basis of the concept of biography focuses on personal and vocational commitment, and intrinsic rewards (Ganz, 2000). Biographical experience is considered the primary source of socialization, cultural perspectives, and motivating interests, encompassing a wide range of demographic factors that bring diversity and its associated benefits (namely opportunities for serendipity and innovation.) Biography is the experiential basis for personal interests and skills, and includes the individual characteristics typically referred to as

“demographics,” e.g., race, gender, educational background, age, etc. In the case studies, aspects of biography that were emphasized most often by interviewees included educational background, age, and community membership. For example, the GSP organizers referred to their participants as “older adults,” and specified that they meant people over the age of 40.

7.4.2 Personal Interests

Motivation, individual goals, and personal values play a major role in participation and outcomes for individuals; these concepts are the basis of the personal interests concept. Motivation, in particular, is of much interest to researchers and project managers (e.g., Raddick et al., 2010; Nov, Arazy, & Anderson, 2011), as it is seen as a leverage point for increasing contribution as well as understanding why people will essentially work for free. Motivations are neither singular nor static, but rather a collection of reasons to participate that change with time and experience (Rotman et al., 2012). Most motivations described by interviewees (across projects) were intrinsic, based on personal interests and values, as well as community-based social rewards that deepen these intrinsic motivations. This is entirely congruent with the concept of biography, but adequately distinct that it merits separate consideration. Although demographics and skills will vary among volunteers involved in different projects, both practical reports and academic theory suggest a number of common motivators for volunteerism, which may have differential effects on individual experiences and performance (Lawrence, 2006; Pearce, 1993; Cnaan & Cascio, 1999).

The concept in the theoretical framework, however, is not motivation, but personal interests. To consider motivation alone would be an oversimplification of the factors that drive individual participation, and though frequently mentioned by interviewees, motivation was not the only variable they described as responsible for initial and ongoing participation. Personal goals and values are both distinct from but related to motivations. For example, a personal goal of making a contribution to science, in alignment with personal values of sup-

porting conservation, can motivate participation (at least temporarily) by individuals whose interest in the actual participation tasks is limited. An organizer for the GSP explained that their participants expressed interests based on personal affinity for an organism (bees) as well as altruistic goals for scientific contribution: “We love bees, and we love what you guys do, and it’s really important” (Bombus, 13081–13216).

7.4.3 Skills

Skills, knowledge, and expertise are individual assets that become project inputs when they are applied to project processes. As noted previously, the Strategic Process Model conceptualizes skills under the label repertoires, which is congruent with the definition used here. The concept of skills is particularly meaningful in the context of citizen science, as participants’ skills are often cited as an important influence on project design, participation, and contributions. Empirically, this relationship is most evident in projects that leverage the range of participant skills in different roles according to expertise and local knowledge, such as eBird’s network of volunteer data reviewers. One of the more interesting aspects of citizen science is the set of mechanisms that are used to accommodate a broad range of participant skills (Wiggins et al., 2011), which connects the personal attribute to design and participation processes.

7.4.4 Networks

The term networks, as used here, refers to the social network or web of personal contacts each individual creates throughout his or her lifetime. These networks are a unique resource that individuals bring to their participation in a project, including both strong and weak ties (Granovetter, 1973). They provide projects access to a broader range of intellectual, human, and material resources, particularly through the engagement of diverse participants. For example, the eBird project leaders mentioned that their personal networks became an

asset for volunteer recruitment, “playing on contacts that I’ve made over my years birding, trying to talk people into helping us” (Stercorarius, 18845–18956).

Ganz notes that sociocultural networks are sources of ideas, mechanisms for recruitment, sources of social capital, and incubators of new collective identities (Ganz, 2000). Networks are inextricably intertwined with communities, but distinct in their egocentricities. The combination of overlapping individual networks yields a broader social network, which typically includes community connections as well as a host of additional personal relationships. In addition, communities are typically oriented around a shared primary goal, interest, or practice, while social networks are formed around individuals.

7.5 Processes

In an IMOI model, the inputs are understood to influence the effectiveness of work groups through two sets of moderators, processes and states. Processes are the dynamic interactions among group members leading to outputs (Table 7.6). Organizational theorists have observed, “organizations do not have mechanisms separate from individuals to set goals, process information or perceive the environment. People do these things,” (Daft & Weick, 1984, p.285). In this theoretical framework, however, all processes, save participation, are conceptualized at the project level because they are established as collective processes that involve both organizers and contributors. Participation, on the other hand, is represented at the individual level because it is carried out by the individual participants who contribute to the larger project. Most citizen science projects are not designed specifically for group participation, although most protocols do allow for it, and others explicitly support classroom participation. Understanding the common processes among citizen science projects is the first step in designing technological and social arrangements that support intellectual production and innovation.

Concept	Level	Brief Description	Origin
Science	Project	Research procedures	Initial framework
Design	Project	Initial & ongoing project decisions	Data
Organizing	Project	Project & volunteer management	Initial framework
Participation	Individual	Taking part in project activities	Initial framework

Table 7.6: Summary of processes from the theoretical framework.

7.5.1 Science

Scientific research processes are an essential part of the framework; as with scientific interests, if there are few or no science-related processes occurring, it is hard to make the argument that a project is actually citizen science. While it may not be fully applicable to every project, most scientific research follows a general hypothesis-testing model, commonly known as The Scientific Method, with the implicit assumption that there is only one acceptable process for generating scientific knowledge. In recognition that there are alternate scientific research paradigms which may be relevant to citizen science projects, “science processes” is used here as a more general term. This distinction is important because, as one of the eBird organizers reflected, “citizen science will never really replace experimental types of research” (Dendroica, 29491–29647). eBird also provides a counter-example to the hypothesis-driven research paradigm, as it is not designed to address specific hypotheses.

The cases selected for this study all share the common contributory model discussed in Chapter II, in which participants assist in data collection but are not involved in other steps of the research process. The science processes mentioned most often by interviewees included design of protocols for participation, data validation or verification, data management, and analysis. For example, a researcher described how qualities of the eBird data had changed the science processes related to data analysis: “every step of this process that we do on our data processing model, we’ve had to change, because of the size and scope of that data” (Passeirina, 45651–46456). Science processes were most frequently connected to scientific interests and scientific knowledge outcomes across the cases, both of which are logical and expected

relationships. Other relationships that were repeatedly highlighted included links to design (particularly for protocols), contributions (a required input to analysis), and participation (with respect to modifications to conventional scientific processes.)

7.5.2 Design

Design is both a process and an artifact, a verb and a noun. In the context of the current work, it is best understood from the standpoint of design decisions. These decisions mark the ongoing design processes in a project and often create artifacts that indicate these choices, such as the Mountain Watch data sheets that progressively changed, showing the evolution of the project design. Interviewees readily discussed the evolution of their projects with respect to design decisions. Aside from the initial formulation of a project concept, however, interviewees rarely differentiated design processes from organizing processes because design decisions are often made in the context of managing the evolution of an ongoing project. Therefore, while making the initial design decisions is seen as a specific process in starting a project, subsequent instances of design decisions tended to be tightly coupled with organizing processes. This type of overlap in processes was not unique to the relationship between design and organizing.

Early versions of the theoretical framework focused on specific aspects of design (task design, technology design, research design), but empirical evidence suggested that they are too tightly interwoven to be meaningfully separated. Nonetheless, citizen science research designs and protocols must reflect careful consideration of job design and task design (Cohn, 2008; Trumbull et al., 2000). Organizational design theories link individual-level inputs and outputs (motivation and performance, similar to personal interests and contributions) to task design (Ilgen & Hollenbeck, 1991), as do theories of volunteerism (Pearce, 1993; Wilson, 2000). An example of the match of personal interests to technology design was identified as a key factor in project success by an eBird organizer : “what’s really driving

eBird [contributions] is the fact that it does a lot of the things that birders want to do” (Dendroica, 19808–20365.) In the case studies, project organizers demonstrated an intuitive grasp of these relationships.

7.5.3 Organizing

Participation and organizing processes are two faces of the same coin. The essential difference is between formal leadership, staff, or organizer roles and contributor, participant, or volunteer roles. While these roles may overlap to some extent (e.g., eBird project leaders are also contributors), there is little evidence as yet to suggest that citizen science projects can be successful without some minimal hierarchical structure to provide leadership and fulfill organizing duties.

Operationally, organizing processes encompass the areas of project and volunteer management. It is possible for a self-organizing citizen science project to emerge, which might substantially change the ways that project and volunteer management are approached; however, this study focuses on projects that are centrally coordinated by an organization or professional researcher. While the term “organizing” might imply initial start-up activities, much like “design,” a variety of organizing processes are ongoing throughout the lifespan of a project. Organizing involves coordination, communication, and other types of articulation work, which is the work required to support core tasks. These activities turn technologies and resources drawn from organizations into facilitative opportunity structures, creating new opportunities for members of the public to engage in scientific research.

Communication is important to both organizing and participation and involves both organizers and participants. Communication is included as a part of organizing processes because it was a critical aspect of the work of organizers, which was the focus for this study. Communication in citizen science projects can include one-to-one and one-to-many relationships, a wide variety of technologies, and numerous genres of communication. Each of the

cases in the study employed different communication strategies, with these choices often based on the constraints and affordances of the project inputs rather than the strategies the organizers felt would be optimal. For example, Mountain Watch organizers were able to use print materials and daily presentations by hut staff to communicate with participants about the project. The GSP relied primarily on an electronic newsletter and web forums, while eBird organizers employed listservs, direct email communication, and articles posted on the website.

The tasks and duties involved in organizing a citizen science project are diverse, and the specifics depend on numerous aspects of project design. A unique aspect of the citizen science context is the applicability of volunteer management processes often associated with nonprofit management, e.g., recruitment, selection, orientation, training, supervision, evaluation, recognition, and retention of volunteers (Pearce, 1993). As a Mountain Watch organizer explained, the day-to-day management of the project included:

Making sure that I have an intern that's making kits, it's basically calling on us to print and put together the kits, and get those distributed to the different locations. Writing the newsletter, overseeing the outreach...and getting that all happening. Promoting it, even within our own facilities, reminding people that it's Mountain Watch time. So yeah, it's just making it happen. (Geum, 33981–34767)

As multiple organizers confessed, these practical details were a much larger effort than they had originally anticipated and they found few resources to help them get started. Nonetheless, common challenges, opportunities, and strategies for success have emerged in the form of “best practices” for organizers. The analysis in this study seeks to clarify the underlying theoretical foundations for these practical recommendations.

7.5.4 Participation

Literally speaking, participation means “to take part,” implying a relationship between an individual and a larger group. For this reason, participation is represented at the individual

level, and individual contributions from participation activities are aggregated as project outputs. Participation in citizen science projects is a form of collective action, linking these processes intimately with the mobilization processes of organizing.

Participation processes lie at the intersection of science and organizing processes in citizen science. Without adequate participation, the science couldn't happen; without adequate organizing, participation is ineffective at best. Design choices were also critical to participation, a relationship that is exemplified by the carefully constructed protocols that participants follow to generate contributions. The focus on producing science, however, often means that participants' roles are highly constrained due to the necessity of following a protocol to generate scientifically useful data.

In the case study data, numerous factors both supported and discouraged participation. For example, the interaction of participation with technologies was often mentioned by eBird organizers, who reported that data entry was a deterrent to participation, but data displays produced by the system were satisfying and encouraged participation. Participation was a particularly multi-faceted concept that was closely related to nearly every other theoretical framework concept, confirming that it is a definitional aspect of citizen science.

7.6 States

States are dynamic properties of the group that vary as a function of inputs and processes; the emergent states of a system can be thought of as a way to understand the health of the system. Evaluating these states should indicate the degree to which the system is fully functional. Hackman's model of group effectiveness includes the group's continued ability to work together as an output, which speaks to the sustainability of the project's goals and social structure. States such as sustainability and satisfaction could be considered outputs, but these concepts are more indicative of a project's current status than production. In

the theoretical framework for citizen science, the states of satisfaction and commitment are observed at the individual level, while sustainability is at the project level.

Concept	Level	Brief Description	Origin
Sustainability	Project	Ability to continue pursuing goals	Initial framework
Commitment	Individual	Ongoing responsibility or obligation	Initial framework
Satisfaction	Individual	Fulfillment	Initial framework

Table 7.7: Summary of states from the theoretical framework.

Prior research on virtual organizations has identified the importance of interpersonal relationships that affect the sense of group community, and thus its long-term sustainability (Markus et al., 2000). There was little evidence in the empirical data of any sense of project-oriented community (which may be due to sampling), and communities of practice to which a project is related seemed to fulfill this role instead. Likewise, the concept of collective identity from an earlier version of the conceptual framework found little support in the data. This does not mean that collective identity is not a relevant concept, but that there was no substantive evidence for it in the current study, and a participant-focused study might find more evidence of identity-related concepts. Another concern is volunteers' level of commitment to the project and how it influences their task performance (Cnaan & Cascio, 1999); the topic was repeatedly raised in interviews with project organizers. Understanding how these factors affect the social and technological barriers to and enablers of participation is important for effective cyberinfrastructure and project designs.

7.6.1 Sustainability

Project sustainability refers to the ability of the project to continue operating as intended in pursuit of the project goals, typically implying maintenance of resource flows. Sustainability can refer to a number of aspects of project operation, such as resource sustainability and technological sustainability. It was also seen as a measure of project success: "it's maintained in a way that is sustainable and continues on and really takes hold within the

volunteer communities... I would see that as a success” (Geum 2009, 27956–28453).

The cases in this study indicated that sustainability was most strongly related to resources, followed by design and organizing, making the point that sustainability does not happen by accident. Although the interviewees had clearly given much thought to sustainability, it is not apparent that it is an obvious or up-front consideration when projects are getting started. This may be particularly true when the scale of participation is difficult to predict, as scale seems to be a critical factor in assuring sustainability as a project evolves. Including sustainability goals in project design and organizing processes is important to supporting increased scale and ongoing operations, but sustainability can be threatened by lack of stability and reliability of supporting technologies, protocols, and sources of funding.

7.6.2 Commitment

Commitment refers to an individual person undertaking project activities as an ongoing responsibility or obligation, and applies both to contributors and organizers. The notion of commitment was frequently associated with the concept of personal interests and mission in the sense of a shared commitment to project goals. Because participation on the part of contributors is voluntary, participant commitment was viewed primarily as the dedication of the individual to the project rather than, for example, a sense of responsibility due to membership in a related community of practice. Regular, ongoing participation is the commitment that project organizers typically desire from contributors, with emphasis from interviewees on harnessing, reinforcing, and increasing participants’ commitment to the project. An eBird organizer gave an example of an individual who “really likes the looks of the bar graph checklists that are produced, and to do that well, you have to go every day, or almost every day. So he’s turned into a real zealot about it” (Otus, 9428–9683).

The commitment of the organizers to supporting project stakeholders, e.g., by providing rewarding participation experiences or openly available data sets, was generally expressed

indirectly but seems no less important than commitment on the part of participants. LeBuhn reported that feedback from contributors suggested that they found participation in the GSP personally empowering, and said “that’s the thing that sort of maintains my commitment. I’m committed to the research, but I really think that’s important, and a little gift” (LeBuhn 2009, 20675–21147). Commitment may also be related to the pervasive participation pattern in which a small proportion of participants make the majority of the contributions, a pattern which has become increasingly evident with the growth of online contributory communities that generate readily observable evidence of different levels of contribution.

7.6.3 Satisfaction

Satisfaction is clearly related to personal interests, particularly when participation is a leisure activity. While motivation speaks to intent, satisfaction speaks to fulfillment. When satisfaction overlaps with personal interests, there is a strong fit between citizen science projects and individual contributors. Satisfaction is conceptualized at the individual level because that is how it is experienced, although other sources of satisfaction may stem from collective or project outputs.

The concept was an emergent theme in the interviews that arose in connection to existing concepts, such as personal interests, organizing, and technologies. One of the GSP organizers explained that respondents to a participant survey indicated high levels of satisfaction linked to project goals: “They are happy because they loved our mission, they support it, and they feel good about supporting the mission” (Bombus, 16847–17242). Dissatisfied participants (or former participants) would be less likely to respond to such a survey, but this also supports the notion that satisfaction is important to retaining active contributors. Some of the less satisfied minority also made their opinions known, however, and according to organizers the primary complaints focused on usability problems and the elimination of free seeds.

An eBird organizer mentioned the importance of satisfying contributors in multiple ways,

saying “I think that’s one of the other drivers of this, besides providing tools, many of the output mechanisms provide recognition for those birders who actually submit data” (Ceryle, 32419–32586). Satisfaction is important to both participants and organizers; although the primary focus was on participant satisfaction, a few instances of satisfaction on the part of organizers were also notable. In particular, organizers whose roles involved only indirect contact with contributors expressed great satisfaction from seeing comments and feedback from participants.

7.7 Outputs

Outputs are the products of the project’s inputs, processes, and states (Table 7.8). Outputs represent the consequences of a functioning group and are often viewed as a proxy for process effectiveness (which is different from ongoing project health as represented by project states). In observation-based citizen science, the outputs are usually data collection and/or analysis, participant experiences, scientific knowledge, and broader impacts. The outputs of projects are challenging to compare across contexts, as different measures bear different meanings in context. This variability often depends on the project mission as well as design and science processes, among other factors (Lawrence, 2010).

Concept	Brief Description	Origin
Contributions	Task & non-task products	Initial framework
Individual development	Learning & socialization	Shirk et al. (2012)
Scientific knowledge	Scientific findings	Initial framework
Broader impacts	Intended & unintended products beyond scientific findings	Shirk et al. (2012)

Table 7.8: Summary of outputs from the theoretical framework.

An important feature of an IMOI model is that outputs themselves become future inputs to the dynamic processes that can then lead to outputs. For example, contributions of data feed back into continuing science processes to produce scientific knowledge. In the GSP, the initial steps of research design were organized by project leaders, after which participants

collected data, creating (in aggregate) a project output. These data were then used for analysis in the continuation of scientific research procedures, which is an ongoing process.

Hackman (1987) emphasizes a duality in outputs, incorporating both task and non-task outputs in the model of group effectiveness. Contributions and scientific knowledge are task-related outputs, but individual development is a type of non-task output. In the theoretical framework, the concepts of individual development and broader impacts represent the non-task outputs of citizen science.

7.7.1 Contributions

The use and value of contributions is predominantly dependent on aggregation, so the concept is represented at the project level. Contributions are task outputs that can usually be quantified or measured, and citizen science projects will often refer to numbers of observations, participating individuals, volunteer hours, and similar measures (Phillips, Bonney, & Shirk, 2012). For example, eBird organizers mentioned a variety of ways to measure contributions, such as the number of species reported, the number of locations for observations, and rates of contribution.

Contributions represent a wider array of project outputs than just data or analysis, however, even though these are certainly the primary and perhaps most important types of individual contribution to citizen science. Additional contributions include meta-participation tasks, such as data review and validation, answers to questions posed by other participants, and suggestions of features or improvements to technologies. Contributions are more easily quantifiable than most other concepts in the framework, but as previously noted, how they are established and what they mean for a project can vary substantially from one project to another. Although most contributions are generated at the individual level, others can be a team effort, a characteristic pertaining primarily to contributions made by organizers, such as reusable methods, protocols, or systems.

7.7.2 Individual Development

Learning, skill development, and socialization are types of individual development that can directly result from participation in citizen science. A variety of other forms more specific to conservation and ecology have also been identified in the ecology literature (Trumbull et al., 2000; Wilderman, 2004; Overdeest, Orr, & Stepenuck, 2004; Brossard et al., 2005; Phillips et al., 2012). The forms of individual development discussed by interviewees and identified in the literature were often direct and potentially immediate results of participation processes, in addition to resulting from ongoing processes.

Scientific knowledge and broader impacts required aggregated contribution outputs and potentially further processes (i.e., continuation of science processes that are sequentially dependent on contributions.) In contrast, many individuals' personal development opportunities occur *during* the participation processes that resulted in production of individual, unaggregated contributions. Participation was consistently identified by project organizers as the primary source for new skills, learning, expertise, and experiences. This observation was also confirmed through participant observation. For example, one organizer saw eBird data as a prompt for reflection on the part of both contributors and organizers, noting that bird migration visualizations “makes you think about...how are we affecting those areas, and if we have a restoration strategy for the Mississippi, that this has influence on birds that then appear all across the eastern United States” (Columba, 10820–11052). Additional influences on individual development included mission, organizing, personal interests, and commitment.

7.7.3 Scientific Knowledge

The production of scientific knowledge advances our understanding about the world (and universe) around us. Scientific knowledge production is a primary goal of citizen science

projects, and therefore represented at the project level; these outputs follow from scientific interests and science processes. The simplest operationalization of the concept are measurable outputs such as data sets, scholarly publications, and methodological advances. Collectively, the scientific contributions of citizen science are in fact substantial: Dickinson et al. (2010) estimated that over 1,000 peer-reviewed publications and technical reports have been produced from eight large-scale citizen science project data sets, and the availability of new sources of citizen science data continues to grow. Out of the cases in this study, the eBird project has been successful at producing scholarly articles, while the others needed more time to refine protocols and collect longitudinal data in order to answer their research questions.

The case studies also demonstrated that *applied* scientific knowledge is an output of citizen science at least as often as contributions to the scholarly knowledge base. The scientific knowledge gleaned from contributions is frequently used to inform decision-making in a variety of areas such as policy, land management, and conservation actions, which is in fact a primary goal for many citizen science projects (Danielsen et al., 2009; Overdevest & Mayer, 2008). This was particularly important to AMC organizers, who included an opinion question in addition to asking volunteers to evaluate more objective indicators of air quality:

[It] relates to the resource management question, that under the Clean Air Act...air quality is one of the main resource values that is protected. And so we're trying to get a sense of what levels are acceptable, and when is it unacceptable, and how does that translate into managing visibility resources and addressing the pollution problem. (Geum 2009, 3884–4353)

Accordingly, the notion of scientific knowledge used in the framework should be understood to include science outcomes that are not a part of the scholarly record. Applied scientific knowledge outcomes should be understood as the scientific findings themselves, rather than the related decision-making outcomes (which are broader impacts.) This may make it more

difficult to assess and compare project outcomes; for example, comparing the value of a scholarly article to a technical report that informs a conservation intervention or policy determination seems likely to be a fruitless effort.

7.7.4 Broader Impacts

Scientific knowledge, both scholarly and applied, are not the only scientific outcomes of citizen science projects (Brossard et al., 2005). In particular, the notion of broader impacts at the project level encompasses unintended outcomes. The concept also includes the intended broader impacts that researchers must often specify when seeking funding, which go beyond addressing scientific research goals. Applied scientific knowledge can lead to broader impacts by contributing to decision-making that may have far-reaching effects, demonstrating an inherent relationship between these two types of outputs.

Interviewees noted the potential of participation to lead contributors toward taking action on conservation issues to support small-scale collective action, to serve organizational missions by providing a new way to communicate with the public, and to improve availability of information for decision-making beyond the original expectations. An eBird portal organizer gave an example of a disaster response situation that eBird data could have improved:

I remember we had...a gas tank spreading all the gas, the petrol in the sea close to a factory where they were treating this petrol, and destroying a penguin colony. And the local consultants never consider this penguin colony, explaining that he did not know that there was something [there]. Of course, if you have this kind of information freely on the web, nobody can say “I did not know.” If we have information on the web, you have to use it. (Diomedea, 16726–17314)

Another broader impact seen in the case studies was the adoption of citizen science data, tools, protocols, and infrastructure by third parties. The adoption and appropriation of these project products supported and streamlined day-to-day operations and also enabled new partnerships external to the citizen science project.

The literature has also noted outcomes such as improved relationships between institutions and communities (Ballard, Trettevick, & Collins, 2008), and increased likelihood of policy engagement by participants (Overdevest et al., 2004). These outcomes can be challenging to measure, however, or may be adequately unique that it is difficult to demonstrate a bottom-line value that can be used as a justification for funding. They were typically presented by the cases in this study as a sort of added bonus to the project outputs and scientific knowledge outcomes. Focusing more closely on the benefits of these non-scholarly products could reveal a much higher level of effectiveness in citizen science than is currently apparent, particularly for projects whose scientific goals are not focused on publication.

7.8 Theoretical Framework Reflected in the Cases

To better understand the relationships between the concepts from the theoretical framework, concept network diagrams were constructed for each case (Figures 7.4, 7.5, and 7.6.) These diagrams represent the relative frequency with which statements in interviews were coded with two different concepts. The thickness of the lines between concepts indicates the relative frequency with which each pair of concepts co-occurred. The thickness of the outline around each concept indicates the relative frequency with which each code occurred. As in the theoretical framework diagram, rounded boxes represent project-level concepts, and squared boxes represent individual-level concepts.

Notably, *not all of the connections are displayed*: a threshold was set to remove the most infrequent connections (e.g., single instances) in order to highlight those relationships that were most strongly emphasized. The degree of emphasis on particular concepts and relationships were affected in part by the interview protocols and sampling. Interviewees were consistently asked identical questions about technologies, organizational arrangements, participant characteristics, and the evolution of the project. Each interview, however, also

followed the direction of the interviewee’s interests and involvement in the case, reflecting his or her personal perspective.

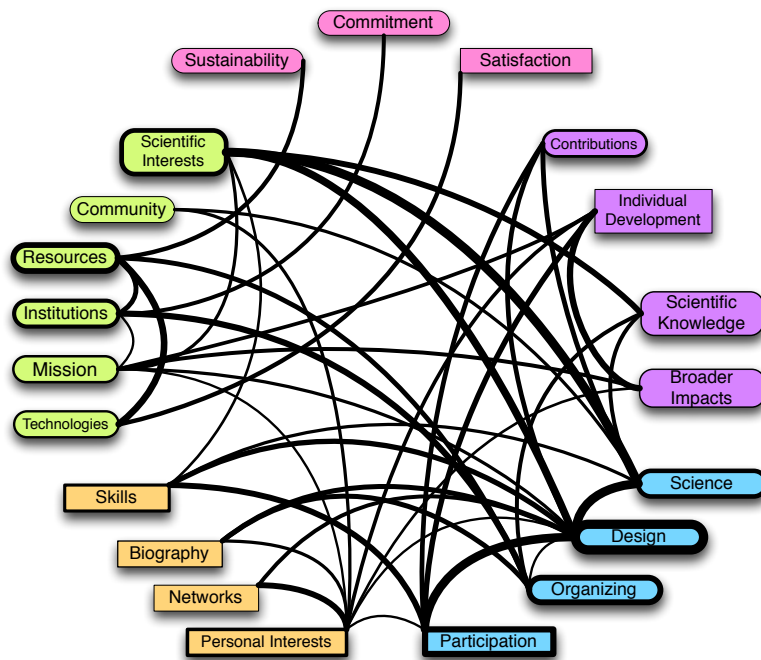


Figure 7.4: Mountain Watch concept relationship diagram.

The three cases showed different relative emphasis on different concepts and relationships between them; this was expected. A number of observations can be drawn from these figures, but the discussion here will focus primarily on the most consistent and strongest relationships. Notably, however, the discussion does not systematically evaluate negative connections; simply because there is no connection shown between a pair of concepts does not mean they are unrelated. The lack of an apparent link between concepts may be due to linguistic, semantic, or analytic artifacts, and indicate potential areas for closer examination. The analysis here focuses on those relationships that were evident in the data to demonstrate the utility of the theoretical framework for surfacing and better understanding the underlying themes in the case studies.

For example, all three projects showed strong emphasis on the relationship between de-

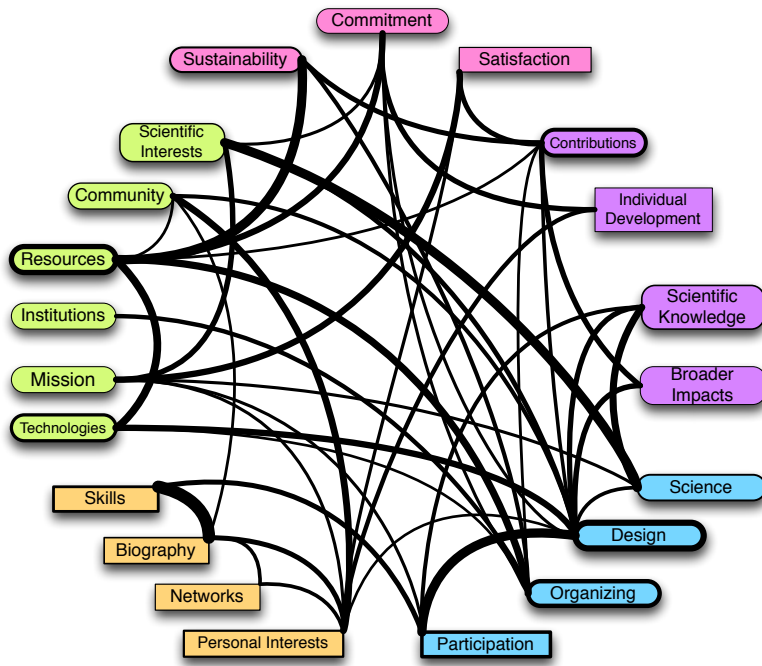


Figure 7.5: The Great Sunflower Project concept relationship diagram.

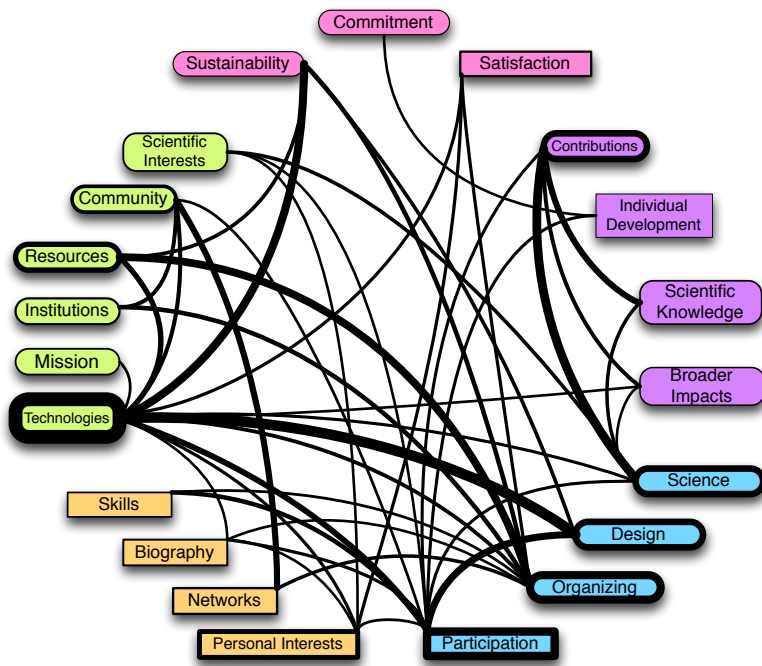


Figure 7.6: eBird concept relationship diagram.

sign and participation, reinforcing the notion that the design of participation processes is a particularly important aspect of all citizen science projects. All three cases demonstrated a heavy reliance on resources to support technologies, emphasized slightly more often by the GSP organizers than in the other cases. Likewise, all three projects noted connections between biography and personal interests, skills and participation processes, science interests and science processes, science processes and scientific knowledge, institutions and organizing, resources and organizing, and resources and sustainability.

The relative strength of the connections also provides insight into the cases. The relationship between institutions and organizing was similarly weighted across all the cases, as was also the case for the relationships between skills and participation processes, biography and personal interests, and participation and design processes. The GSP put substantially higher emphasis on the relationship between resources and sustainability, reflecting the low level of institutional resources available to support the project on a long-term basis.

Strong connections between scientific interests and science processes were evident in Mountain Watch and the GSP, but were much weaker in eBird. The difference could be explained by eBird's lack of focus on a single research goal, as it is instead designed to collect data amenable to a wide variety of scientific research. The lack of a connection between design and science processes for eBird provides further evidence for this explanation.

Several instances where relationships are shared by only two cases are also telling. For example, there was a connection between technologies and design in both the GSP and eBird, but not for Mountain Watch, for which technologies were not central to participation. The concepts of mission and personal interests were connected for Mountain Watch and the GSP, but not eBird. This reflects the eBird organizers' stronger emphasis on self-satisfaction rather than altruistic mission-focused participation, which is related to the project design approach and a topic of later discussion. Along similar lines, a connection between technologies and

design was observed for both the GSP and eBird, in which ICT was a fundamental part of participation. No such relationship was evident for Mountain Watch, in which ICT was much less central.

Interestingly, the connection between community and personal interests were evident in both Mountain Watch and the GSP, but not in eBird, where the personal interests of members in the community of practice were thoroughly integrated into project design. This connection was taken for granted in discussions by the eBird organizers, for whom there was little question about the personal interests of community members. Comparatively, the personal interests of hikers and gardeners were brought up much more frequently by organizers in the other cases, who were less certain about the alignment of the personal interests of community members and the project goals.

Several of the relationships highlighted here are relevant to the topics of the next chapter, in which the synthesis of the theoretical framework and cases focuses on the multiplex relationships between these concepts in the context of the cases.

7.9 Summary

This chapter focused on the theoretical framework developed throughout the course of the research. The discussion of concepts drawn from theoretical models from other disciplines showed the congruence between the citizen science framework and other models, as well as the existence of concepts that bridge these phenomena and fields. It also documented key aspects of the development of the theoretical framework.

Next, the concepts included in the theoretical framework were systematically examined. The relevance of each concept for the phenomenon and rationale for its inclusion in the framework were discussed, with examples from the case studies to support these claims. The chapter concluded with a high-level comparative analysis of the relationships between

the concepts from the theoretical framework that were evident in the case studies. The analysis considered the application of the theoretical framework to empirical study of citizen science, and found logical and expected relationships between concepts that recurred across the cases. There were also notable variations on a case-by-case basis that were reflective of the context of each case study, and a number of the observations briefly highlighted here will be discussed further in the following chapter.

CHAPTER VIII

Cross-Case Analysis

This chapter discusses the findings from cross-case analysis of the case studies, synthesized in an analysis of five thematic topics that relate to theoretical concepts from both the framework and the research questions. The emergent findings from the case studies include five thematic topics that relate to theoretical concepts from both the framework and the research questions.

1. Citizen science project design approaches that favor science versus hobbies for participation design.
2. Project design and organizing implications of engaging communities of practice.
3. Relationships between physical environment, technologies, participant experiences, and data quality.
4. Information technology tradeoffs: helpful for scale and communication, challenging for usability and resources.
5. Resources and sustainability relate to institutions and scale of participation.

8.1 Comparative Analysis of Case Studies

In the previous section, a theoretical framework developed through the research process was presented in detail. The system-level framework provided an evolving lens for examining rich data and highlighted meaningful relationships between important aspects of citizen science when conceived as a virtual organizational structure. It provided a useful tool for

advancing the analysis and interpretation of the data, and can be put to additional uses to support practitioners by providing focus for planning and evaluation.

However, as with any such model, the theoretical framework cannot fully elucidate the nature of the relationships between the concepts it has helped to identify. Many aspects of the concepts and their definitions remain open to interpretation and will benefit from further verification and elaboration in future research. This chapter focuses on discussing emergent themes related to the theoretical framework, which will in turn be used to address the research questions in Chapter IX.

The emergent themes, generated through an iterative inductive process described in Chapter III, can be readily associated with concepts from the theoretical framework (see Table 8.1). The associations demonstrate the relevance of the concepts for understanding the phenomenon explored through the case studies. It also shows the degree of interconnectedness between these concepts in the context of citizen science, although not every connection between the topics and theoretical framework concepts is fully examined in the following sections.

The remainder of this chapter will focus on describing the deeply intertwined, multiplex relationships between these concepts, addressing the research questions through synthesis and grounding in comparisons between the cases. Each of the five themes discussed in this chapter relates to the others, and yet stands alone. Like many other complex sociotechnical systems, a single unifying narrative to neatly combine these diverse facets of the phenomenon would be a fiction.

8.2 Scientific Work Versus Lifestyle Activities

Project design approaches that favor science procedures versus lifestyle (leisure or hobby) practices as the starting point for participation processes are related to the majority of the

Emergent Theme	Individual Inputs	Project Inputs	Processes	States	Outputs
Science versus lifestyle	Personal interests, Biography, Networks, Skills	Scientific interests, Community	Participation, Organizing, Design	Satisfaction, Commitment	Contributions, Individual development, Scientific knowledge
Communities & practices	Personal interests, Biography, Networks, Skills	Scientific interests, Community, Technologies	Participation, Organizing, Design, Science	Satisfaction, Commitment	Contributions, Individual development, Scientific knowledge
Participation & place	Personal interests, Biography, Skills	Scientific interests, Community, Resources, Technologies	Participation, Organizing, Design, Science	Commitment, Sustainability, Satisfaction	Contributions, Scientific knowledge, Broader impacts
ICT affordances & constraints	Personal interests, Skills	Institutions, Resources, Technologies	Participation, Organizing, Design, Science	Satisfaction, Sustainability	Contributions, Scientific knowledge
Organizing	Biography, Networks, Personal interests, Skills	Community, Resources, Institutions, Technologies, Mission	Participation, Organizing, Design	Sustainability	Contributions, Scientific knowledge, Broader impacts

Table 8.1: Association of thematic topics with theoretical framework concepts.

concepts in the theoretical framework. These choices were observed to impact all of the science-related factors in the cases, particularly through the organizing and design processes of developing usable protocols. The contrasting project design approaches are connected to several concepts in the theoretical framework, shown in Table 8.2.

Theme	Individual Inputs	Project Inputs	Processes	States	Outputs
Science versus lifestyle	Personal interests, Biography	Scientific interests, Community	Participation, Organizing, Design	Satisfaction, Commitment	Contributions, Individual development, Scientific knowledge

Table 8.2: Association of project design approaches with theoretical framework concepts.

The typical revision cycle for protocols in the dominant science-oriented project design strategy had implications for resource requirements and the time required to produce scientific outcomes, both of which played into project sustainability. The alignment of scientific interests with personal interests that the project design choices reflected further guided the way participation processes evolved. Some projects offer more options and others narrow the tasks, with meaningful impacts on the scientific outcomes that can be achieved. The

narrowing and expanding of participant roles connects to participant autonomy. This characteristic of participation design affected satisfaction, commitment, and the ways that personal interests and skills could be leveraged to incentivize ongoing participation and increasing contributions. The topics of revision cycles, alignment of interests, and participation design are the focus of the following sections.

8.2.1 Revision Cycles

The typical science-oriented project design adapts conventional science processes by reducing their complexity for public participation. As a result, most such projects experience several revision cycles, major or minor, that can add up to years of delays before scientific outcomes can be achieved. This was not the case for eBird, which based its original protocols on existing community practices.

As a project leader reflected, “We have changed some of the ways they [the protocols] appear or formatting but the definitions are essentially the same” (Pinicola). The early issues related to data quality and usability for scientific research were instead due to which protocols participants selected when reporting observations. When the ways that protocols match birding practices were clarified and emphasis was placed on the more scientific protocols, the volume of observations submitted according to the more scientifically useful protocols increased substantially in a very short period of time. These changes to birding practices did not require participants to change to the way they already performed the fundamental task (observing birds), but instead meant incrementally augmenting those activities with additional information. Providing additional rewards to contributors in the form of new technology features (e.g., yard lists) further incentivized contribution of the most desirable data for research purposes, again without any changes to the protocols themselves.

As mentioned, this is an unusual situation in citizen science. By comparison, the GSP and Mountain Watch projects followed the usual model of adapting a scientific task to a

format suitable and palatable for public participation. In both cases, the necessary cycles of protocol revision slowed the production of scientific outcomes. Piloting the procedures was inadequate for fully identifying and addressing issues with protocols that were identified only when the projects were launched for full-scale participation. Since both projects were also seasonal in nature, each revision cycle took up to a full year to collect the data and analyze quality. Year-long revision cycles add up. When the process of producing a viable protocol to meet scientific goals takes several years, initial funding is quickly exhausted before the project is able to produce results or truly engage participants at full scale. This leaves organizers in a particularly tough position: just when the project is poised to succeed, resources evaporate.

Simply put, making a case for supporting a citizen science project on scientific merits is even more difficult when results have yet to materialize. At the same time, the evolution of the participation protocols and processes seems inevitable for most projects.

8.2.2 Participation Process Evolution

Most citizen science projects, particularly those taking a science-first design approach, require multiple revisions to the scientific protocols that impact the participation processes. The net effects of these changes lead to an evolution in the alignment of scientific and personal interests, as seen in the eBird case: although the birders were always interested in birds, only when the project catered to that interest was the desired scale of participation achieved.

The nature of changes to participation processes, made to support improvements to the scientific outcomes of the project, can also affect the experience of participation, with ramifications on participant satisfaction, commitment, recruitment and retention. These changes either broaden or narrow the task options and roles that contributors can play. Examples of broadening the task options included additional protocols and ways to contribute (such as the yard lists) in eBird and the addition of phenology monitoring in the GSP. Another

broadening of options came from the addition of new monitoring plants in GSP.

Broadened tasks give participants more choice in the way that they participate, which provides them with a different level of autonomy in relation to the project participation expectations. It also enhances project appeal because there are additional personal interests that participation may serve, and can lower barriers to entry. For example, if an individual was already growing rosemary and Purple Coneflowers but not Lemon Queen sunflowers, they could still join the GSP mid-season and contribute data from existing garden plants, rather than having to wait until the following year to plant the one specific species that had been designated for monitoring.

By contrast, a narrowing of the participation tasks occurred in Mountain Watch, and the net effect likely also lowered barriers to entry. Reducing the number of scientific tasks expected of hikers who have little domain expertise and less direct affinity for alpine wildflowers (compared to the trails and mountains) simplified participation. Reducing task complexity may have made it easier for less confident contributors to decide to participate. At the same time, Mountain Watch organizers did not remove all participant autonomy: although specific locations were marked to simplify the monitoring process, there was also space on the data sheet to accommodate making observations at other locations, as well as a general data sheet that would allow entirely self-guided location selection. This choice prevented stifling the interests of more experienced, advanced, or enthusiastic contributors, continuing to provide them greater autonomy with respect to the location choice for participation, even if these data are less useful than those collected at permanent plots.

Broadening and narrowing the participation tasks therefore indicated common probable consequences across cases based on potential to expand participation. The convergent results of divergent strategies suggests that rather than over-simplifying or complicating participation processes, each project worked toward a middle path that permitted just enough choice

without overwhelming contributors, although the organizers may not have recognized it at the time. Across the cases, this evolution also had similar benefits for scientific outcomes.

Influences on scientific outcomes reported by organizers included better data quality and higher volume of contributions through repeated participation and increased participant satisfaction. The task expansions in both eBird and the Great Sunflower Project addressed known participant interests that contributors had expressed to organizers. The Great Sunflower Project organizers broadened the participation tasks in responses to popular demand by adding popular garden species which best fit the scientific goals. For eBird organizers, the patch and yard list features were developed to elicit more contributions of the most valuable data for research by providing rewarding feedback. As members of the recreational birding community, the project leaders were able to identify this opportunity based on their own interests, but had also previously received participant feedback requesting features that would incentivize collecting lists from the same location, which further confirmed the utility of the approach.

Both Mountain Watch and the GSP had repeated revision cycles, discussed above, that also affected the broadening and narrowing of the projects. Identifying the data quality problems in Mountain Watch contributed to the narrowing of participation tasks to better accommodate skill deficits. Interestingly, direct observation and attempts to replicate the participant data collection process were necessary to identify the challenges volunteers encountered with the participation process. For the Great Sunflower Project, direct observation of participants was not feasible due to the geographic scope of the project. The challenges of spatially distributed work highlight contrasting affordances and constraints with respect to the observability of participant behavior and the potential for resulting changes to participation processes to better accommodate participant autonomy, interests, and skills.

In eBird and the GSP, the choices to modify participation tasks were also linked to

individual-level states of satisfaction and commitment. Adding desirable features that broadened ways of doing the core tasks in eBird, for example, increased the commitment of the yard list and patch list competitors. Adding desired monitoring species for the GSP offered additional satisfaction to the existing participants. In both cases, the changes led directly to increased contribution volumes and repeated sampling at set locations, which also yielded improvements to the overall data set. For Mountain Watch, the narrowing of the task may have resulted in greater participant satisfaction due to reduction of uncertainty, but the potential effect was not confirmed.

Across the cases, these changes also leveraged personal interests and the known contributor skills to incentivize ongoing participation, repeated participation, and increased volume of contributions. By adjusting task options to meet a wider range of interests and skills, these practices reflected job design principles related to participant autonomy (Ilgen & Hollenbeck, 1991), which would be expected to enhance commitment and satisfaction (Millette & Gagné, 2008).

Offering greater autonomy is in direct conflict with the standards of conventional science, which emphasizes uniform protocols and carefully controlled processes, removing as much autonomy as possible in data collection. In the cases in the study, participant autonomy was accommodated in a number of ways, such as the inclusion of optional protocols and additional choices for monitoring species (the GSP), choices of locations for participation (Mountain Watch, eBird, and the GSP to a lesser extent), frequency of participation (eBird, the GSP, Mountain Watch), and choices of observation protocols, timing, and duration of participation (eBird). eBird's participation processes, which mirror community practices more closely than scientific research processes, allowed participants substantially more autonomy to fit participation into their lifestyle than most citizen science projects. This tight alignment between scientific data collection processes and existing community practices contributed

substantially to eBird's success in generating scientific outcomes.

Comparatively, protocols derived from conventional science practices tend to narrow the required and available options for participants over time, as seen in Mountain Watch. This trend, which often places greater emphasis on data quality at the cost of participant satisfaction, may also lead to reduced retention in the long term as participants find little room for expansion of their participation. Careful design of self-reinforcing participation experiences, however, can make a substantial impact on the appeal of ongoing participation and was observed to lead to increased commitment for some contributors.

8.2.3 Design of Self-Reinforcing Experiences

To state the obvious, making participation fun incentivizes ongoing contribution, supports participant commitment and satisfaction, and eases recruitment and retention of contributors. In two of the cases, organizers explicitly mentioned the moderating desire to avoid beating people over the head with the task, making project engagement something that volunteers want to do.

Accomplishing this goal is easier when the task can be based on an existing community practice. The example from eBird shows a type of participation in which the scientific task is secondary to a pleasure activity and a minor extension of existing practices. In science-first project designs, however, it is more difficult to create alignments between participation tasks and community practices. Most conventional science tasks bear no semblance to hobby pursuits and can be somewhat awkward when packaged as such, so convincing volunteers to undertake new tasks may be more challenging. In particular, organizers may have to work harder to convince would-be participants that they are adequately capable of the task, which is potentially less interesting to them and may not fit as easily into day-to-day activities.

An example of a way to modify a protocol from a science-first to a community-first design was mentioned by a Mountain Watch organizer. Instead of asking volunteers to do

a variety of identification tasks in the field, participants could instead make photographs of the monitoring plants at the designated sites. The photos could later be uploaded to the Mountain Watch site, allowing examination and verification of the phenophases for data entry, either by the photo contributors or by others, e.g., through online games. While there are a number of problems that would have to be solved to make such an approach feasible, it represents a participation process that is closer to a pleasurable, easy task related to what hikers already do.

In all the cases, project organizers acknowledged that participation should ideally produce mutual benefits to be truly sustainable. Accomplishing this goal was a matter of making technology development investments for the eBird project, as the interests of birders were well known by project organizers. As previously mentioned in the discussion of the GSP, however, learning what contributors would like to get out of participation is not always so straightforward. At a general level, prior research has addressed the issue to some extent with studies of volunteer motivations (e.g., Raddick et al., 2010; Nov et al., 2011; Rotman et al., 2012).

Another job design principle that can be leveraged to support self-reinforcing participation is creating room for advancement. Most citizen science projects provide only one role for volunteers. Creating multi-level participation structures can allow advancement (e.g., eBird's hotspot reviewers and network of editors) and harness existing or growing expertise. Although additional organizing effort is needed, these opportunities provide new challenges that help maintain interest, and reward contributors with acknowledgement of their contributions and skills.

8.3 Communities and Practices

Compared to the prior literature on online communities, the cases in this study showed a pattern of citizen science projects that connected to existing communities rather than creating their own, blurring the lines with respect to community membership and identity. Whether projects work with existing communities versus building new communities influences project development, as well as the perspective on community as an input versus an output.

The communities participating in a citizen science project and their existing practices highlight relationships between several theoretical concepts (Table 8.3). Integrating existing communities into citizen science projects can impact participation processes and protocols, particularly with respect to science-oriented versus lifestyle-oriented project design approaches. Existing community practices can also influence design of technologies, participant expectations, the available skills, and participant training needs. Community structure can moderate the value of partnerships with existing communities. These themes are discussed below, with a focus on the issues of community integration versus creation, project protocols and community practices, and the effects of community structure on volunteer management.

Theme	Individual Inputs	Project Inputs	Processes	States	Outputs
Communities & practices	Personal interests, Biography, Networks, Skills	Scientific interests, Community, Technologies	Participation, Organizing, Design, Science	Satisfaction, Commitment	Contributions, Individual development, Scientific knowledge

Table 8.3: Association of communities and practices with theoretical framework concepts.

8.3.1 Community Integration Versus Community Creation

An intuitive insight expressed by numerous citizen science project managers is that it is harder to build a new community than to partner with an existing community. The cases

in this study worked to partner with existing communities. The strategy of developing a new community seems more likely to occur primarily in citizen science projects undertaken in solely online environments or focusing on scientific inquiries that do not have logical links to established communities of practice.

Integrating a citizen science project with an existing community provides numerous opportunities. It can help project organizers find ways to appeal to community interests and present project features with the right message for the community. At the same time, the eBird case included a contrary example drawn from the cultural differences between North American and French birding groups. While the basic community practices of bird observation and data sharing were the same, the ways that data were shared were different due primarily to the French tradition of using these data for financial support of conservation activities.

Developing a new community instead would imply developing a sense of shared interests, creating new community practices, and establishing venues for communication and social interaction, rather than simply piggybacking on existing structures. Although these tasks represent substantial challenges for developing social infrastructures to support project participation, they are also likely to provide new opportunities exclusive to developing a community from the ground up. For example, a project that essentially creates a new social group of individuals who otherwise have no affiliations could find a greater level of commitment from participants because it provides a venue for aggregating individuals with shared interests that previously had no “home.” The potential of a project to create new communities can be seen in other contributory online communities such as Flickr, where numerous localized photo groups developed out of the new support for identifying and organizing fellow photographers in a geographic region (Cox, Clough, & Marlow, 2008).

8.3.2 Aligning Protocols and Community Practices

Another important aspect of integrating a community into a citizen science project requires aligning the participation processes with community practices. The concept of a community of practice, as introduced in Chapter II, defines these groups according to shared practices, such as those involved in birding. Aligning community and scientific practices is a matter of degrees. Most projects seem to find it difficult to adapt existing hobbies into scientific protocols. Whether a project is designed from a science-first or a community-first approach, a balance of scientific rigor and appeal for participants has to be achieved. This is always a challenge.

The alignment of scientific and personal interests was particularly notable in the eBird case, where protocols were modest adaptations of the existing community practices. Increasingly scientific forms of participation were offered and encouraged as a natural progression of commitment to the project goals, as well as providing data-oriented rewards for participation. In addition, eBird's data outputs supported existing community practices like reporting sightings to email listservs and friendly competition for local records. A distinct benefit of the approach is that birding tends to be a lifestyle hobby, with avid birders making regular observations and trips to view birds. This makes it easy for participants to incorporate eBird participation into their existing habits.

In the GSP and Mountain Watch cases, the participation tasks were less familiar for contributors. The protocols for observation in each of these cases were based on traditional scientific processes. While unfamiliar in general, the GSP tasks were far simpler and required no special skills, making the project appear accessible to a wider range of potential contributors with less expertise. Nonetheless, the tasks bear little similarity to existing practices, meaning that participants had to explicitly make an effort to include the tasks in their gardening routines; establishing new habits around unfamiliar tasks can be difficult. The

Mountain Watch project design, however, required a number of tasks that were difficult for hikers and had little precedent in standard community practices. As a result, the participation protocols were simplified over time as the challenges in the tasks became evident to organizers.

The eBird participants' enthusiasm for birds further represented alignment of community interests with scientific goals, which could be seen in the GSP as well. As previously discussed, birds, bees, and sunflowers are all considered charismatic species. Many people are excited about and attracted to them. Among the citizen science practitioner community more broadly, there is a general sense that only charismatic species are able to draw this type of following. Several counterexamples do exist, however, in projects focusing on slimy, unattractive organisms such as algae ¹ and American eels ².

A notable departure from the organism-centric appeal of citizen science for communities of practice was seen in the Mountain Watch project. Hikers care more about the landscape than the particular organisms. Although they clearly find the wildflowers attractive, evident from the annual calls received by AMC inquiring after blooming times, doing anything more than taking a photo and momentarily admiring their beauty—as part of a broader landscape—is not necessarily a common practice among hikers. The challenging nature of the focal practice of hiking rough terrain also makes it much more difficult for hikers to adopt a regular habit of stopping to monitor the flowers.

Another particularly interesting aspect of community practices was the individual changes to established community practices seen among devoted eBirders. As they became better able to see the alignment of their interests with the scientific interests, some birders were willing to undertake a slightly more complex task set to improve the scientific value of their observations. While the more rigorous protocols are considered “harder” than the usual

¹<http://www.backyardbiofuels.org/participate.html>

²<http://www.dec.ny.gov/lands/49580.html>

community practices, it is worth acknowledging that by comparison to most citizen science participation protocols, they are simple. Further, the protocols are elaborations of existing practices, rather than entirely new routines.

Maximizing observation value in eBird requires complete checklists, counts of birds, and effort information, all of which are relatively easy tasks. They can also be adopted as a slow progression of intensity of participation. The flexibility of the protocol options makes it easier to convince contributors to gradually modify their birding practices as they become comfortable with each task added to their established habits. For example, after contributing to eBird for some time, a birder might begin collecting effort information during their birding trips, the basic prerequisite to making the data scientifically useful. After awhile, jotting down effort information becomes habit, requiring no second thought. Then the individual might start making complete checklists, encouraged by the data displays that show species distribution for a location throughout the year. Along the way, he or she might turn to counting and estimating population numbers, which is an interesting challenge that is repaid with reports that enable new comparisons. The optional evolution of existing habits to include new elements makes it substantially easier for individuals to progress along the participation learning curve at their own pace.

Finally, both eBird and the GSP realized that accommodating community practices with respect to data submissions was important. In both cases, all data were accepted, even those which were not considered currently useful for research. There were several reasons for this choice: it demonstrated and reinforced the organizers' messages that all contributions are valued; it satisfied participants who like to keep records; and it acknowledges the potential future value of these data. Doing so comes at minimal cost to the science processes if database structures are properly constructed. It is a normal part of scientific practice to verify and clean data, which is the main drawback to accepting these contributions. It could

be construed, however, as a subtle encouragement for submitting substandard data, but organizer messages to the contrary were quite clear in both cases.

Accepting all data aligns with community practices by permitting individuals to participate to the degree that they are comfortable. In the case of eBird, it acknowledges standard birding practices that organizers hope to influence by rewarding improved quality. For the GSP, it aligns with gardeners' usual practices: few individuals plant sunflowers alone, and few gardeners attend carefully to only one species. It is much more likely that an enthusiastic gardener would want to spend equal time admiring their rosemary plants, for example, in addition to Lemon Queens. The alignment of community and scientific practices through the mechanisms described here are best practices in volunteer management, discussed further next.

8.3.3 Community Structure and Volunteer Management

Highlighted most strongly in eBird, community structures were valuable for participant recruitment in citizen science. The utility of community structures is true for most volunteer projects, and recruitment of contributors is a core volunteer management task in nearly all such contexts. This theme was echoed across all three cases, as several modes of accessing networks and communities provided multiple avenues to expanding participation.

In some cases, such as the GSP and Mountain Watch, project organizers were able to tap into multiple communities to support project recruitment. The GSP organizers focused on developing connections to different communities, including gardening groups, adult education organizations, and youth groups like 4H. For Mountain Watch, the focus was on recruiting additional contributions through partner organizations, including parallel organizations like the Adirondack Mountain Club. While these organizations have formalized institutional structures, they are also made up of communities of practice.

Recruiting through established organizations is an approach similar to the strategy es-

poused in social movements literature. By taking advantage of indigenous organizations with existing community structures, citizen science partnerships can gain more value for the investment in recruitment and volunteer management effort. eBird also used this approach, reaching out through a wide variety of partner organizations. These included numerous localized groups around the world, which was an integral part of the strategy for globalization with continued staffing at the same levels that previously supported North American participation alone.

As already discussed, indigenous organizations provide special resources that support participation. In particular, communication structures are among the more useful features when it comes to recruitment. Embedded communication structures may also explain the projects' continued successes despite providing relatively low levels of social support for participation. Individual contributors' membership in indigenous organizations and communities of practice could instead provide social support and reinforcement for contribution through existing community values.

The example of using eBird's emailed checklists to forward to listservs shows the utility of appropriating community infrastructures and practices. The organizers did not attempt to replace existing communication networks or practices, which would likely have met with resistance or disinterest. Instead, they created tools that align with those practices, both supporting participant satisfaction and providing an additional mechanism for advertising the project to a much wider community. Something as simple as the footer reading "generated by eBird 2.0" attached to each checklist forwarded to a local email group provides visibility that would otherwise be difficult for project organizers to achieve. These characteristics of communities and networks suggests that there is less need for technologies to support social interaction when a project is situated within a community structure that provides existing communication venues.

Another tactic for supporting volunteer management through access to communities was hiring respected non-scientists from the community of participants to manage citizen science projects. Hiring community members was enormously successful for eBird, as it provided deeper access to community structures that could support volunteer management, was a form of endorsement of the project by respected community leaders, and brought extensive insight into ways to align scientific and community practices. The GSP unknowingly emulated this example when LeBuhn hired Bombus, who not only had several skill sets that were needed for project development, but was also a locally respected community member. The benefits of the staffing choice were not yet fully evident at the time of data collection for this study, but several new partnerships suggested that the strategy would provide new options for project development. A notable difference between the hiring choices for the GSP and eBird is that Bombus is known locally, while the eBird project leaders had attained broader community recognition through high visibility activities like leading birding tours and editing birding magazines.

In both cases, the importance of hiring a non-scientist was emphasized. This was considered key. Even though at least one was a scientist by training, the three eBird project leaders were recreational birders first, and their understanding of and commitment to the recreational birding community likewise came first. LeBuhn also clearly identified the value of including a non-scientist as an organizer when she remarked,

I thought that bringing someone who's more of a passionate gardener to complement my science skills, he's much more like the participants than I am, so I thought he would bring a really strong perspective that would complement my skill set. I mean, I love gardening, but I'm a scientific gardener...I'm not a garden show person. (LeBuhn 2010, 5079–5640)

A similar attempt to tap into non-scientist perspectives was evident in the Mountain Watch case when an organizer recruited non-botanist friends to participate under observation. In-

corporating a non-scientist hiker as a project organizer was also achieved through partnership with the Education department: “When we first came at this as scientists, we definitely had a different approach, and they [the Education department] helped us move it more towards something that is more palatable for a citizen” (Geum 2010, 14773–15213).

Through a combination of approaches, each of the citizen science projects in this study leveraged community practices, structures, and shared interests to support project goals. The next section focuses on the role of place in supporting participation and its effects on the scientific outcomes of the projects.

8.4 Place and Participation

Place seems to be taken for granted in project design except as a constraint, and even then it is often treated as a foregone conclusion. The challenge of defining place is aptly summarized by Gieryn (2002):

“Place” is not easily defined, but might usefully be conceptualized as having three necessary and sufficient features: (a) Place is a unique spot in the universe, a geographic location of elastic bounds; (b) place has a physicality, and its material form variably combines natural environment and built architecture; (c) place holds meanings and value, and it is the object of labile and contested narrations and imaginations. (p.113)

As this quote suggests, place plays different roles in different citizen science projects. Place profoundly influences the way the science is constructed and what can be produced by a project.

Theme	Individual Inputs	Project Inputs	Processes	States	Outputs
Participation & place	Personal interests, Biography, Skills	Scientific interests, Community, Resources, Technologies	Participation, Organizing, Design, Science	Commitment, Sustainability, Satisfaction	Contributions, Scientific knowledge, Broader impacts

Table 8.4: Association of participation and place with theoretical framework concepts.

The natural environment poses different types of limitations and benefits that were particularly relevant to the cases in this study. The places in which participation was carried out impacted technologies, participation processes, who could participate, participant retention, scope and scale of participation, and resulting contributions, along with other conceptual relationships shown in Table 8.4. Scientific processes and outcomes are influenced by these factors, manifest in such details as the relative importance of the degree of location precision and repeated data collection from single locations. Place is related to participants' personal interests, commitment, and autonomy. These factors can be linked to technologies, place attachment, and ground truthing; these topics are the focus of the following discussion.

8.4.1 Technologies and Places

Relatively little research in organizational studies, information systems, and related disciplines has taken on deeper considerations of place with respect to information technologies, with the notable exception of Harrison and Dourish (1996). In another study of contextualized use of technologies, the way documents were utilized in a healthcare setting was seen as a representation of places and related interdependencies of individuals, times, and spaces (Oesterlund, 2008). In Oesterlund's study, documents were found to act as portable places representing the interdependencies of subject, object, time, and space. The documents served as both communication and coordination tools, and also played a role in socialization and learning, much like the role of data sheets in the citizen science projects discussed in this study.

For the cases reported in this study, place had a substantial impact on the way technologies were used: paper technologies were used in the field, but digital technologies were used largely after the primary participation task had occurred. The preference for material technologies in the field was clearly due to the influence of the physical environments where participation occurred, places that are often prohibitive of effective information technology use in the field

due to a wide variety of factors. In their study of technologies and place, Harrison and Dourish (1996) described the relationship between space and place as the difference between opportunity and the understood reality, emphasizing that space is a location but place is the locus of action. The specifics of participation activities in citizen science impact the choice of technologies for different tasks; out of all possible options, paper was consistently preferred for use in the field³. The affordances of paper were a practical reason for the use of data sheets due to the characteristics of the places where data were collected.

In addition, however, some citizen science project leaders also collect paper data sheets as a backup to digital records (as reported by respondents to the survey mentioned in Chapter III). Sometimes these material reference points provided additional information about the data, even when they were duplicated electronically. The relationship between these paper and electronic records further emphasized the material and cognitive divide between field-based participation and data entry tasks. As eBird project leaders noted, an observation submitted to the database is the unique intersection of a person, a bird, a time, and a place. The observations themselves, whether recorded in material or digital form, represented these inherent interdependencies in observation-based scientific data collection, just as in the prior study by Oesterlund (2008). In eBird, there were no project-specific field data entry sheets, an acknowledgment of the existing ingrained styles of field-based record making in the birding community. For the GSP and Mountain Watch, however, the data entry sheets were another tool to remind participants of the details to record and the process of doing so. The data sheets and online data entry interfaces impose a structure on participation and their repetition supports independent learning of the protocols, for which the data sheets are a material extension.

Another notable relationship between time, place, and documentary forms is related to

³This was changing as of early 2012, when third parties developed tools to permit eBird-specific data entry on mobile devices with applications designed around contextualized birding practices

the ways that scientific knowledge is created and organized, e.g., in classification systems (Bowker & Star, 2000; Bowker, 2005). Temporal and spatial markers are a way of organizing knowledge, and a way of transforming data into knowledge. This is self-evident in the influence of temporality and location for the scientific value of the data produced by the citizen science projects in the study, particularly with respect to the importance of repeated observations taken in the same place over time. In eBird, recording the date, time, and location with more precision improved the quality of the data for use in scientific analyses, and verifying that multiple observers were present further supports the perception of truth attached to these observations. The precision of place was even more important for Mountain Watch, leading to substantial revisions of the participation protocol, and by extension, to the material technology of the data sheets in which place was inscribed and described in detail.

In reference to classification as a form of knowledge organization connected to place, the unique taxonomy that eBird implemented to make its reporting system more robust to mis-identification was a particularly interesting way to handle the relationships between individuals, places, and species. The taxonomy allows birders to track subspecies and morphs (variations in plumage) that mark primarily geographic variations in species. These details are often uninteresting to researchers who are likely to be interested only in the fact that a Dark-eyed junco (*Junco hyemalis*) was observed; it is the birders who care whether the junco subspecies was Slate-colored, Gray-headed, Oregon, Pink-sided, or White-winged. Places are remembered by birders according to the species that were observed there; the first place a species was seen is often jotted in the margins of field guides, making the annotated books a valued reminder of place.

Being able to record these experiences, regardless of scientific value, is a point of individual satisfaction that was important for the eBird system to support. Some of the more

expert individuals contributing data put great stock in these variations, as some birders also collect subspecies in the same way they collect species for their life lists. Having seen every subspecies of Dark-eyed junco can represent a substantial accomplishment among avid birders. Observing them in person requires travel to new places, and the memories of places are inextricably entwined with the species observed there. The relationship of individuals to place through experiences is also reflected in the concept of place attachment, discussed next.

8.4.2 Place Attachment

The cases in this study highlighted the advantages and disadvantages of place attachment, a concept from environmental psychology that refers to the bond between people and places. Two facets of place attachment, both of which are relevant to these cases, are place dependence (functional attachment) and place identity (emotional attachment) (Williams & Vaske, 2002). These characteristics are discussed for each of the projects, as they had different impacts on the way the scientific aspects of the project were designed and the participation experiences of contributors.

The White Mountains provide an excellent example of the interrelationship between place dependence and place identity. With respect to place dependence, the goals of Mountain Watch can only be achieved with data collected in alpine zones, a rarified habitat that brings with it numerous constraints on participation. As project organizers quickly learned, hikers on vacation are a difficult audience to recruit for scientific work. The challenges of identifying locations in this environment also created problems for acquiring scientifically useful data, as more precision was needed than most participants could readily provide. The weather conditions were often difficult and limited participation in several ways. All of these functional limitations influenced the design of the protocol from a scientific standpoint.

At the same time, place identity counters these constraints with the strength of contribu-

tors' emotional response to the place. The White Mountains are nothing short of spectacular. The landscape itself is exceptionally beautiful, with expansive vistas full of natural wonders that pique curiosity and evoke a strong emotional response.

Although the infrequency of return visits by project participants was among the challenges organizers named with respect to project sustainability, they also noted that the place has a powerful pull on many people: “A lot of folks that come to these mountains do have a real connection to these mountains. They come back here regularly, whether it’s yearly or bringing their kids back to the same places they went to [as children]” (Ledum, 3861–4196). Without the nightly reminders from hut croos, however, past contributors might not remember to pick up a monitoring kit when they return the next year. Although the emotional aspects of place attachment are largely positive in connotation, this suggests that there are also drawbacks for citizen science projects in such special places.

By contrast, the GSP observations nearly always take place in a personal garden, although shared gardens are permitted and even promoted. Place dependence for the GSP is based on where one can grow a plant for monitoring bees, whether that is a potted plant on an apartment balcony, a thin swath of soil in a narrow city front yard, or a meticulously maintained backyard gardenscape in a lush rural setting. This feature of the project means that nearly anyone can participate, not just those with the physical constitution to climb mountains. Plants grow only where planted, so it also means that repeated contributions for any particular participant will come from the same location, which has substantial benefits for the research that the data can support.

These gardens are nearly always at participants' homes, which creates a very different relationship of the participant to place from the Mountain Watch experience. The ordinary front yard landscaping is also transformed by participation. It changes from an everyday taken-for-granted, and therefore nearly unseen space, into a place where important scientific

tasks can be accomplished. New meaning is attached to chores like thinning seedlings, incessant weeding, driving away predators, and staking up exceedingly large plants that otherwise topple under their own weight. Instead of being done solely for the sake of a pretty yard, although pleasant on its own, gardening also becomes a way to contribute to science.

Still, the mundanity of the location can easily wear away an initial sense of novelty. The bright blooms of sunflowers can be transformed from an exciting potential to contribute to something bigger into a nagging reminder of promises broken for contributors who fail to follow through on a commitment to participate. The everyday location and the intentional visibility of garden flowers have a special attribute can be both positive and negative: they serve as a constant reminder. As with Mountain Watch, the consistency of the locations for participation in the GSP presents limitations as well as benefits from the perspective of encouraging participation.

eBird demonstrates a very different model with respect to the role of place: place dependence does not limit participation. The variety of places for participation both reflects the normal community practices and helps keep participants interested. Supporting participant autonomy almost by accident, the project can take place anywhere that wild birds are seen. Even the most urban environments contain birds, albeit often only a limited number of “boring” species.

Removing geographic limitations on the places where participation can occur dramatically widened the scope and scale of data that can be contributed to eBird. According to organizers, this was a key factor in recruiting some individuals who were uninterested in contributing data when participation was constrained by continental boundaries. One partner project organizer did mention that for some monitoring locations, contributors must be able-bodied enough to get into and out of cars, but at the project level, people with physical

limitations can still report birds viewed through windows of cars, houses, and nursing homes. The primary constraint with respect to place dependence is the contributors' willingness to report data at finer spatial scales, listing birds that occur at a local park rather than somewhere within a county. At the same time, the system effectively incentivizes more detailed reporting behavior by providing location-based reports that have strong appeal for birders.

From a birder's perspective, any place is made more interesting by the presence of birds. Birders are notorious for hanging out at wastewater treatment plants—hardly an attractive place by most measures—in order to spot species that frequent these locations to feed. Birders return to the same places year after year, for example, to see migrating warblers decorate otherwise nondescript places with their bright colors. The autonomy of eBirders in choosing places to record data has positive effects on who participates, how often, how repeatedly, and the overall volume of data that are collected.

More importantly for the scientific interest of the eBird organizers, birders will assiduously record the birds in some locations with surprising frequency. Some contributors keep daily yard lists, and some locations attract so many birders that they are also surveyed daily. Such regularly collected location-specific data can be hard to come by in the citizen science world. One of the most substantial challenges for the research goals of eBird data users, however, also stems from the freedom of participants to choose their own observation locations: geographic bias. Project organizers have tried without success for some time to find mechanisms that will effectively convince birders to visit “non-birdy” places, as the natural inclination is to take a bird walk in a place where interesting species are most likely to be seen. In the meantime, and likely for the foreseeable future, statistical modeling techniques have been developed to address these geographic biases in large-scale analysis of the eBird data.

In addition to these nuances of place attachment, the role of place with respect to truth in the scientific enterprise is also a meaningful consideration, discussed next.

8.4.3 Ground Truthing

Ground truthing is a practice that refers to data verification, typically in reference to remote sensor data (Dictionary.com:2012b, 2012b). The ground truth of remotely gathered measurements is verified in the field for calibration and interpretation; for example, the process that Mountain Watch employed for assessing data quality was a form of ground truthing for citizen science data rather than satellite imagery. In scientific vernacular, the phrase can also refer to presumed objective reality, as opposed to reported expectations thereof.

Ground truthing has interesting implications for the relationship of place to data quality, acceptability, and verification methods:

The field site must be made into a place that persuades: Scientific authors situate themselves there and present a relationship between knower and place that would enhance insight, objectivity, accuracy and trust. The place itself must be variously assigned qualities that carry epistemic freight—exotic or close, typical or unique, pristine or instructively invaded, empty of people or full of them. (Gieryn 2002, p. 118)

This quote highlights the way that place is relevant to establishing the truth of scientific findings. Not only must place be persuasive, but in citizen science, the individuals in that place must also be persuasive, adding another level of complexity to establishing truth. When the research is conducted by professional researchers, both of these factors seem to fade into the background. The discussion of place becomes a question of generalizability that is addressed through methodological means that are often poorly suited to citizen science data collection, which instead tends to be more opportunistic with respect to data collection locations.

As the Mountain Watch case showed, inability to reproduce the species identifications due to the imprecision of location description was a symptom of a larger concern. Modifying

the observation protocols to enable some level of expert verification essentially supported the ability to ground truth the participant data, triangulating hikers' observations against those of trained naturalists. Such careful ground truthing is hardly a scalable approach; it was functional in Mountain Watch because of the particular resources available to organizers and the geographic scope of the project.

Each of these data quality and validation methods hinges on the specifics of place. The acceptability of data from places where participants cannot be observed in the process of contributing is frequently questioned by professional scientists. Respondents to the survey discussed in Chapter III indicated that the inability to evaluate individual skill through direct in-person interactions was a substantial source of concern for data quality. The uneasiness with respect to data quality is rooted in the intersection of ground truthing and virtuality: researchers cannot be in the same place and time as the observers.

Data quality concerns point to the inherently distributed nature of participation in citizen science as a fundamental "flaw" from a scientific perspective. At the same time, distributed collaboration among professionals is becoming increasingly common, so the issue of collaborator expertise is also at the heart of the concern over data quality. This is in fact the primary apprehension voiced by most skeptics of citizen science, but the fact that it arises from issues of place and virtuality that interfere with ground truthing goes unmentioned.

The degree of virtuality in terms of the spatiotemporal spread of contributors suggests that distributed citizen science, while addressing the constraints of place, simultaneously brings scrutiny to the fact that these issues are overcome through volunteer participation. In these cases, the conservatism of traditional scientific practices is linked to place in addition to the expertise concern that is often cited. While many citizen science organizers are aware that contributors may have far more expertise than assumed, the lack of colocation and direct contact makes the degree of individual expertise difficult to evaluate.

The notion of ground truthing also has a deeper relationship to truth in science. Professional researchers have addressed most issues around the role of place in establishing scientific truth, but citizen science leads to renewed concerns about the representativeness of sampling, the precision of location data, and other factors that are related to expertise but prompted primarily by virtuality. The fundamental enabling factor that permits this approach to distributed data collection and processing is ICT, which in addition to the benefits of permitting these new forms of collaboration at more and more places, also introduce constraints; these are discussed next.

8.5 ICT Constraints and Affordances

The degree of reliance on information and communication technologies is another factor that influences who can participate in citizen science projects. ICT can play an important role in citizen science, and surfaces other dependencies and relationships (see Table 8.5), in addition to its own constraints and affordances.

Theme	Individual Inputs	Project Inputs	Processes	States	Outputs
ICT affordances & constraints	Personal interests, Skills	Institutions, Resources, Technologies	Participation, Organizing, Design, Science	Satisfaction, Sustainability	Contributions, Scientific knowledge

Table 8.5: Association of ICT with theoretical framework concepts.

ICT afford increased scale of participation and easy modes of communication, leading to numerous benefits for many citizen science projects. Digital technologies also come with issues of cost, customization, and usability, particularly with respect to data entry. Data entry was a major hurdle for every project, but participant characteristics and feedback also played roles in enabling and incentivizing an otherwise dull task. As a result of the dependence of ICT design and functionality on resource flows, many citizen science projects find ICT development and support tasks overwhelming. When they provide the foundation for virtual participation, however, ICT can have a substantial influence on data quality,

scalability of participation, scope of research, and organizing processes. These factors in turn affect the science processes and scientific outcomes. This section focuses on the affordances and constraints of ICT for citizen science, highlighting the themes that were emphasized by project organizers: scale, communication, usability, and costs.

8.5.1 Affordances: Scale and Communication

The obvious affordances of ICT are the ability to increase the scale and scope of participation in citizen science. The communicative functions that ICT serves by providing greater visibility, modes of technology-mediated interaction, and new ways to provide feedback to contributors are the main affordances responsible for enabling increased scale and scope. The cases in this study employed multiple approaches to using technologies to increase the scale of participation. For both eBird and the GSP, online data entry is a fundamental aspect of the project that permits large-scale organization. eBird also worked with other organizations to provide portals, kiosks, and data exchange tools that further spread the project's reach. Mountain Watch also implemented online data entry to extend its potential contributor base, but because it has never been as central to participation as in the other cases, it has seen relatively low adoption due to staffing limitations on recruitment across a wider geographic range.

There were also unexpected affordances of ICT that worked to organizers' benefit, such as the time-lapse videos produced by Mountain Watch, which took advantage of plantcam outputs to augment training materials. Both eBird and the GSP noted the potential of online data entry to improve data quality by controlling input options and permit automatic error-checking. Form field controls are an easily overlooked feature of web-based data submission that has substantial benefits for citizen science, in addition to reducing data entry and validation effort on the part of organizers. Numerous other broader impacts of the technologies were seen in eBird, where the system has become a form of infrastructure for other

organizations and projects. Likewise, the GSP's observation protocol has been adopted by other pollinator-related projects, an outcome that would not have come about without easy access to and visibility of the project's website.

The GSP also demonstrated how low-cost technologies can have high associated costs, but would permit running a large-scale project from a home instead of an organizational setting. This is good news for projects that do not have institutional resources to provide technological infrastructures. For nonprofit and voluntary organizations, however, relying on volunteers to provide technology support is a strategy that is often doomed to failure because it is rarely sustainable. Therefore the potential benefits of low-cost technologies come with the caveat that they often require customization and suffer usability problems that can present serious challenges to project development and growth.

At the same time, ICT can provide a rewarding experience to contributors. The eBird project, which has made substantial investments in developing new features that cater to birders, is a good example of the potential value of ICT-based feedback. In particular, several interviewees noted that visualizations are very engaging to individuals—not just contributors, but also data users and potential participants—and help them see the bigger picture. As the GSP project showed, however, it is not always clear what types of data access and visualizations are appropriate, valued, or relevant to contributors. In such cases, investment in costly technology development for data visualization and reporting may be better undertaken gradually in response to interests as directly expressed by contributors. Involving community members as organizers and contributors of feedback on project development are ways that project organizers sought to address these challenges.

8.5.2 Constraints: Usability and Development Costs

Two of the major constraints related to ICT that were persistent across the cases were issues with data entry and the cost of technology development. A related underlying theme

was that of usability, which is a potentially problematic issue for scientific outcomes.

As mentioned above, low-cost technologies such as open source software often present usability issues and need customization to be minimally functional for citizen science project needs. Concerns over usability are especially relevant for older users who often make up the bulk of project contributors. Organizers consistently hoped that technologies would help engage younger and different audiences, but aside from faster adoption by younger birders in eBird, there is little evidence from the cases suggesting that the intuition is true in practice. This may be more attributable to recruitment strategies than the actual potential of technologies to draw the interest of new audiences, but the excitement over the inherent appeal of ICT may well be unwarranted. Younger generations may simply expect higher technology production quality (a concern expressed by one of the eBird technical staff) and find novelty only in the most cutting-edge developments, many of which are out of reach for citizen science projects.

Fortunately, prior research can provide substantial insight into technology design standards to ameliorate the usability concerns that are particularly important for older adults. The GSP project leaders had identified research on usability for older adults that they found particularly revelatory with respect to prioritizing usability needs. Their experiences reinforce the often-repeated refrain that knowing the audience for a citizen science project is not just helpful, but necessary, as the usability issues identified by the participants themselves were reportedly exacerbated by age-based considerations. For large-scale projects in particular, initial investments in usability and motivating features may pay off in reduced future costs.

The problems with usability experienced by project organizers were largely centered around data entry forms and processes. These are readily solved by engaging professionals in technology development and customization, an expense that should clearly be considered

necessary rather than optional due to the potential impacts on participation and data quality. Motivating data entry, however, was another matter entirely. The eBird project managed to overcome the detraction of data entry by providing immediate system-based feedback and visualizations that are considered particularly rewarding and interesting by contributors. As previously mentioned, though, the form of feedback that may motivate different audiences is not always clear.

Numerous ideas to motivate and reward data entry can be drawn from other research communities, such as human-computer interaction and game design, as well as entirely technology-mediated citizen science projects such as the Zooniverse projects. At the moment, these strategies are rarely implemented by observation-based citizen science projects. It is likely that some of these mechanisms are either unknown to many organizers of such projects, or are considered too expensive to develop, low priority in comparison to functionality and usability needs, or possibly even trivial. The argument could also be made, however, that the lack of scalable feedback and reward systems will eventually undermine project growth and participation sustainability.

Another related consideration is that immediate feedback may help reassure contributors that their performance is adequate and that they are performing core observation tasks well, even if they find the technologies intimidating or challenging to use. Such reinforcement can convince individuals who might otherwise quickly abandon the project, to persist long enough to get over the learning curve that eBird organizers specifically mentioned as a barrier to contributor commitment to ongoing participation.

In all of the cases in this study, ICT development and management was an important organizing task, but only one of many that will be discussed next.

8.6 Organizing

In the cases examined for this study, organizing was most substantially influenced by available resources, institutional influences, project scale, and task complexity, though additional relationships became apparent (Table 8.6). Numerous aspects of organizing were discussed in the individual case descriptions, but two themes bear further consideration given their repetition across the cases.

The biggest constraint on project growth and development was limited organizer time and attention, which was in turn related to resource flows. Organizers consistently reported underestimating the effort required for basic project development and implementation, much less expansion. Communication seemed to suffer most from organizer constraints. Communication problems can have cascading effects on participation since communication can be instrumental in recruitment, retention, and improving data quality. It may therefore be beneficial to allocate and adjust staffing based on the scale of participation and complexity of protocol. By extension, meta-contributors who help support the work of other volunteers can further improve scalability, but also require organizer time and effort to coordinate.

Theme	Individual Inputs	Project Inputs	Processes	States	Outputs
Organizing	Biography, Networks, Personal interests, Skills	Community, Resources, Institutions, Technologies, Mission	Participation, Organizing, Design	Sustainability	Contributions, Scientific knowledge, Broader impacts

Table 8.6: Association of organizing with theoretical framework concepts.

Assets like supportive organizational relationships and infrastructure shape project development and impact project sustainability. Sustainability influences organizing and project development (and therefore scientific outcomes) in ways that can influence altruistic participation. Sustainability strategies vary, and are an important consideration during the initial stages of project development. Managing project scale and sustainability were recurring

themes across the cases.

8.6.1 Organizer-Contributor Ratio

One of the most promising aspects of citizen science is the ability to harness the energy and attention of a large group of people to accomplish tasks that cannot be handled by a smaller number of people. When recruiting efforts are successful, large-scale projects can lead to very high numbers of contributors for the number of organizers. Each of the cases in the study encountered challenges related to the ratio. The more contributors and fewer organizers, the more challenge there was with maintaining adequate communication and feedback, particularly for the GSP. At the same time, the approximate ratios of active contributors to organizers for each of the cases were surprisingly similar, suggesting a potential natural limit to the number of participants that a project can support with a given amount of organizer resources.

The types of participation activities and skills of contributors are also linked to the issue of participant support. When activities are simple and skill levels are high, a small number of organizers can more readily handle a very large number of participants. When either of those conditions are not true, more organizer time is required to adequately support participation. The interaction of skills and task complexity suggests that the number of participants that organizers can support is not a fixed ratio. In the eBird case, the challenge was met through partnerships that offloaded some of the organizing effort to partner project leaders, e.g., portal managers. This strategy took advantage of the strengths of indigenous organizations, as previously discussed, which is a more scalable approach to managing large numbers of contributors.

Without enough organizer time to match to task complexity and skill levels of the contributors, additional challenges emerge. Managing data quality becomes more difficult, particularly if expert review is part of the process, as it does not scale well. Answering questions and

addressing participants' problems—managing online accounts was mentioned as a continual task for organizers—becomes an overwhelming task when a single organizer is faced with 70,000 individuals signing up for the project. Even managing volunteers to manage other volunteers in a meta-contribution structure becomes too difficult under these circumstances, as many meta-contribution tasks are not as well defined or simple as the core contribution tasks.

This appears to be a challenge with no single solution for citizen science project organizers. Successfully managing a large number of participants seems to lead to increasing levels of participation in a positive feedback loop. Increasing participation is generally desirable for citizen science projects that can benefit from larger scales of contribution, but means the demand on organizers is magnified with continuing success.

There is a practical limit on the efficiencies that organizers can achieve with respect to managing a large number of volunteers. It seems that once the upper limit on the organizer-contributor ratio is reached, which likely varies by project, additional organizer resources are required to maintain the status quo or to continue to expand. This suggests that projects should include contingency plans for staffing increases or the addition of meta-contribution opportunities in project sustainability strategies to avoid becoming a “tragedy of success.” Planning for project sustainability is an important consideration that is discussed next.

8.6.2 Institutions and Sustainability

The interconnection of organizing, institutions, resources, and sustainability were evident in the concept relationship diagrams discussed in Chapter VII, which showed consistent connections between institutions and organizing, resources and organizing, and resources and sustainability. Organizational settings provided a safety net and resources that supported project sustainability during early development, which helped both eBird and Mountain Watch. Working with partner organizations was a strategy that reduced the overall organizer

effort while extending project reach for eBird and Mountain Watch, and seemed promising for the GSP as well. These efforts at developing a broader network of organizers to support project activities required attention and effort, but had potential to substantially increase the project capacity for growth and evolution. At the same time, the eBird case highlighted ways that partnerships could influence goals and outcomes in citizen science. Changes brought about by partnerships are often to the benefit of all involved, but can pose their own set of challenges as organizers seek to support a wider range of partner activities. Crafting mutually beneficial institutional arrangements is key to producing sustainable partnerships, an outcome that can also require institutional support to achieve.

The eBird case showed that a diverse portfolio of resource streams was needed to keep the project sustainable for long-term operation. Indeed, while the difficulty of getting large grants frustrated some organizers, it also forced them to take more entrepreneurial approaches to project sustainability. An entrepreneurial funding strategy may be more sustainable but means that organizers can end up spending more time chasing dollars than ensuring project success or research outcomes. When it comes to sustainability planning, there is also a major learning curve on the part of scientists who are used to the grant funding paradigm rather than selling products, soliciting donations, or negotiating sponsorships. This may be a tractable issue for organizers who are able to rally institutional resources in some settings, but could become problematic for academics who are expected to focus every moment of attention on research and publication.

Developing sustainability plans that relied less on institutional resources in the long run also appeared to be important for both eBird and the GSP. As both the GSP and Mountain Watch organizers noted, start-up funding is easier to obtain than ongoing funding, but these cases also showed that more start-up funding to cover longer development times may be needed. The NSF Informal Science Education (ISE) funding program recognizes this need

with a series of grant awards that build upon one another, starting with a planning grant that allows organizers to work out practical details before moving on to full-scale implementation. Notably, however, not all citizen science projects are focused on ISE as a project goal.

For the GSP, LeBuhn was unsuccessful with grants for NSF funding focused primarily on scientific knowledge production, suggesting that citizen science has been institutionally pigeonholed as an outreach activity. Unfortunately, LeBuhn's experience demonstrates the reluctance of the scientific establishment to risk funding an innovative approach to large-scale research. While there are outreach and education benefits to the GSP, there was not enough ISE focus given the scientific goals, so the projects could not be funded by either a scientifically-focused grant program nor the ISE grant program. Part of the challenge could be the seemingly slow rate at which most citizen science projects produce results. NSF or similar funding could help improve pace of producing results, however, if funding could be applied to producing appropriate technologies to support ongoing participation and good data collection.

An overarching concern for citizen science is the importance of early planning to support long-term sustainability, as single grants are not sustainable and strong initial funding does not guarantee continued support. Many citizen science projects focus on long-term monitoring or other goals that cannot be achieved in just a few years. Organizers do not always plan ahead for sustainability, however, or their plans may be predicated on conditions that turn out to be inapplicable to their situation. For example, including contingency plans for large-scale participation seems particularly appropriate in light of the experiences of the GSP organizers. It seems unlikely that project organizers currently include this scenario in their initial project planning processes. As practitioners learn and share more about their experiences organizing citizen science projects, these and other issues for sustainability will doubtless continue to emerge. Documenting practices that lead to both success and failure

will be an important ongoing activity for the broader citizen science practitioner community as it proceeds to grow.

8.7 Implications

This section summarizes several implications of the findings. These include the implications of case sampling on the findings themselves. Additional discussion focuses on the implications of the findings for scientific outcomes in citizen science and scientific collaboration more broadly.

8.7.1 Implications of Case Sampling

As discussed in Chapter III, the cases selected for this study are representative of the broader population in most respects, but vary widely in the resources available to the project organizers, including projects from opposing ends of the general distribution. In particular, including the eBird case in this study shaped the findings in several ways. Its commonalities with the Great Sunflower Project (e.g., with respect to hiring non-scientist community-oriented organizers) served to highlight additional prospective best practices. It also demonstrated that the issues that all three cases encountered in the development of supporting technologies can be overcome; eBird successfully addressed challenges associated with usability and motivating data entry through participant-focused ICT design and feedback of data outputs.

Had eBird been omitted as a case in this study, however, the findings would likely have concluded that citizen science is an unsustainable and inefficient approach to generating scientific knowledge. At the same time, both Mountain Watch and the GSP showed clear signs that this approach does have potential for much more substantive outcomes, which have been shown in other cases that were not included in the study. Instead, the results showed that with adequate resources and resourcefulness, citizen science projects can reach

a tipping point in which the value substantially outweighs the investment.

Without the eBird case, there would have been little evidence of integration of community practices with scientific protocols, as the project design approach for the other two cases followed a more typical science-first path. There also would have been very little support for the conclusion that the experience of citizen science can be self-reinforcing in any respect other than altruism. The findings would have assumed that extensive revision cycles are always necessary for developing an effective participation protocol, which may be the case for most projects, but should not be as universal a situation as new projects begin adopting the protocols of existing projects. The outlook for benefits from ICT would have been far less optimistic, because many small citizen science projects are limited by underdeveloped ICTs that pose more constraints and generate fewer benefits in terms of promoting participant feedback and satisfying user experiences. There would have been less richness but little change to the findings regarding place or organizing, aside from the aforementioned point regarding project sustainability.

By including the eBird case, which is an exemplar in project practices, the findings were more balanced. Instead of a relatively discouraging set of findings, the evolution of smaller projects along a similar trajectory to a very large and successful project holds out the potential that citizen science can in fact generate the scientific outcomes that most organizers hope to achieve. The variation across these cases helped illuminate the ways in which the differences between the projects influenced outcomes, thereby providing potential explanations for recurring issues facing smaller citizen science projects and suggesting avenues for improving project performance.

8.7.2 Implications for Scientific Outcomes

The scientific community continues to question the value of data generated by citizen science projects, and whether they truly generate good science. In the cases in this study,

Mountain Watch and the GSP were motivated by formal research questions, while eBird took a different approach. The GSP's research questions focused on understanding bee visitation rates across different habitats in a study designed to investigate the larger issues of pollinator service, which are a key part of ecological processes. Mountain Watch had two research questions centered on mountain plant phenology that would identify indicators of climate change through the study of phenology. eBird took a different approach, as previously discussed, and collected data on bird abundance and distribution that could be used to answer a variety of scientific research questions and were suitable for integration with complementary data sets.

The eBird project has unquestionably demonstrated that high quality scientific outcomes can be achieved through citizen science. Among over 80 scholarly publications directly stemming from eBird data, numerous articles discussed the project itself. There are many others, however, that demonstrate quality scientific research outcomes, including articles in highly regarded journals such as *PLoS Biology*, *Ecological Applications*, *Evolution*, *PLoS ONE*, and *Proceedings of the National Academy of Sciences*, among many others.

The scientific outcomes of smaller citizen science projects are not always evident or impressive. Not all citizen science projects are as strongly focused on generating scientific knowledge as the cases presented here, in which scientific knowledge outcomes were a top priority, and scientific interests and science processes played significant roles in shaping the projects. In this study, both Mountain Watch and the GSP had yet to publish scientific findings due to ongoing modifications to address data quality issues. However, both projects had recently achieved a level of data quality that would permit them to produce rigorous scientific findings that address research questions focused on important environmental issues.

In the GSP, the primary limitation was the availability of the researcher to complete the work. This situation could have been improved by the addition of collaborators from the

science community (e.g., graduate students), or the reduction of project management duties. While the case study reported that there were issues with some data in terms of subjective reporting of location types, these issues were being addressed and the scientific data itself was considered adequately sound for use in peer-reviewed publications. For this project, evidence for the scientific merit of the work was just around the corner.

For Mountain Watch, however, it was not as clear that quality science would emerge from the citizen science portion of the project. The complementary data sources will make it possible for the long-term outcomes of the effort to produce valid research findings to address the research goals. The ability of hikers to generate data adequate for rigorous scientific application was still in question. Substantial effort was invested to identify the problems creating shortcomings in the scientific data. New modifications tested in 2010 and shown effective for improving data quality were implemented in 2011, and organizers expected that these changes would make it possible to collect data adequate for generating rigorous scientific research outcomes. Additional time would be required, however, due to the nature of the topic of study: phenology research typically requires long-term data sets. Another notable consideration is that AMC's mission focuses on conservation, not research; research is a means to conservation outcomes and therefore academic publications may not be the first or most prevalent scientific knowledge products for this project or others like it.

On the whole, this is a troublesome trend. It appears that high quality scientific outcomes are in fact possible—good science is being accomplished. The pace at which it is achieved, and the relative value of the research outcomes in comparison to the resources invested in citizen science is more questionable. It also seems quite likely that many smaller projects will never achieve their initial goals with respect to scientific outcomes, likely due to mismatches of protocols to participants and shortages of human resources with respect to organizers and scientists. Even when adequate resources are available, the time required to set up a project

capable of producing outcomes like those seen in eBird is typically far greater than expected by organizers, leading to concerns about project sustainability.

Nonetheless, there is evidence that citizen science can be good science and produce valuable contributions to scientific knowledge. Improving these outcomes is a clear direction for further development in citizen science.

8.7.3 Implications for Scientific Collaboration

As discussed in Chapter II, citizen science projects show a surprising degree of similarity to scientific cyberinfrastructure projects. The degree of specialization and routinization of scientific work in these two contexts suggests that the structures of many large-scale scientific collaborations may be less distinct from citizen science than previously expected, particularly in terms of the functional versus intellectual roles of contributors. The parallels between these phenomena suggest that perhaps there can only be so many intellectual contributors in collaborative work.

Issues of coordination costs and task interdependency are explicitly addressed in the participation structures of citizen science. Implicitly, this practice suggests that involving too many people in the intellectual work may be an ineffective (or inefficient) work structure, as opposed to engaging larger numbers of contributors in the functional supporting work that is frequently a feature of large-scale scientific collaboration. Prior research has noted that task interdependency varies widely across scientific collaboration structures. Findings across studies suggest that low interdependency increases productivity and collaboration success, particularly in distribution collaboration environments (Walsh & Maloney, 2007; Cummings & Kiesler, 2005; Perlow, 1999; Olson & Teasley, 1996).

In the context of Wikipedia, research has shown that a larger number of editors (collaborators) improves article quality only when the work is coordinated implicitly, with a few editors doing most of the edits and others supporting their work (Kittur & Kraut, 2008).

If we view the organizers of a project as taking a parallel role to doing the larger part of Wikipedia edits, based on the notion that these efforts represent the majority of the intellectual work, this distributed collaboration pattern is remarkably similar to that of large-scale citizen science and scientific cyberinfrastructure projects. In citizen science, minimizing coordination costs through a task structure that relies on pooled interdependence has been very productive, particularly in the eBird case. The similarity between Wikipedia and citizen science, and concurrent contrast to other forms of distributed scientific collaboration, suggests a need for further study of task interdependence in distributed collaboration.

Related questions pertain to how many levels of hierarchy are involved in each context, and what that implies for the necessary management structure, as well as the qualitative differences between contributions across roles within the collaboration. For example, as the case studies demonstrated, citizen science typically has a relatively shallow hierarchy of roles with most participants doing the same task, which reduces coordination costs. This is similar to the bureaucratic and semi-bureaucratic structures of scientific collaboration identified by Chompalov et al. (2002), discussed in Chapter II. Future work could therefore investigate whether these work and social structures currently exist or might be effective in scientific collaborations that do not involve members of the public.

In addition, this study found that there is a need for methods of evaluating scientific knowledge production beyond publication and citation counts, particularly with respect to the applied scientific knowledge outcomes that are important in many citizen science projects. Sonnenwald (2007) reinforces the importance of these “other” outcomes in conventional scientific collaboration. As new ways of demonstrating scholarly impact emerge, so do calls for new modes of evaluation for scientific work, especially for large-scale research initiatives (Trochim et al., 2008). The repetition of this theme suggests that further investigating ways to evaluate a broader range of outcomes in citizen science may shed new light on additional

outcomes of value in other scientific collaboration environments.

8.8 Summary

Each of the themes discussed in this chapter can be related to the concepts from the research questions (see Table 8.7) as well as the concepts from the theoretical framework. For each of the topics, relationships to the research questions were observed for every key concept. This too demonstrates the complexity of citizen science projects when viewed as sociotechnical systems and virtual organizations.

Theme	Virtuality	Technologies	Organizing	Participation	Scientific Outcomes
Science versus lifestyle	distributed collaboration	incentives	sustainability, revisions	recruitment, commitment, satisfaction	volume, quality
Communities & practices	networks	incentives	recruitment, revisions	recruitment, commitment, satisfaction	volume, quality
Participation & place	expertise, truth	affordances, constraints	opportunities, constraints	recruitment, commitment, individual development	volume, quality, precision
ICT affordances & constraints	scale	usability	design, costs	incentives, data entry	volume, quality
Organizing	scale	management, design	institutions, sustainability, funding	communication, skills	sustainability, volume

Table 8.7: Association of thematic topics with research question concepts.

When it comes to citizen science project design approaches, two dominant models were observed: science-first and lifestyle-first participation protocols. The science-first approach is dominant, but seems to require additional revision cycles that typically result in reduction of participant autonomy, although the GSP provided a counterexample in which participation options were broadened. The lifestyle-centric approach is less common but meant that the eBird project avoided revision cycles, could offer substantially more autonomy in participation, and was able to easily create increasingly self-reinforcing participation experiences. In each case, however, changes to protocols and modifications that expanded or narrowed participation options all supported participant satisfaction and improved outcomes for the

scientific work.

The role of community in citizen science was seen more of an input to the project than an output, meaning that existing communities were leveraged rather than new ones created. Community practices and project practices intersected whenever an existing community was integrated into a project, which seems to be a dominant model in observational citizen science because of the obvious benefits for volunteer recruitment and knowledge of audience interests. The converse approach would be to build a community around a citizen science project, which is generally seen as more challenging and was not the strategy chosen by any of the cases in the study, but may be necessary in other contexts. Working with an existing community indicates a need to align participation protocols with established community practices. Project leaders also considered what activities and outputs community members would find interesting as well as what skills and background they could bring to the project. In addition, existing community structures helped ease volunteer management duties by providing easier access to networks of individuals for recruitment, as well as established modes of communication and sometimes leadership that effectively extended the project staffing resources.

The discussion of the role of place and its influence on participation noted that place is often a taken-for-granted constraint (and opportunity) when it comes to citizen science project design. Technologies in particular were strongly influenced by the environmental conditions in which participation was carried out. These factors also influenced science and participation processes and several aspects of participation. The analysis extended these observations by considering the materiality of technologies used in the field and the way that spatial and temporal markers are used to transform data into knowledge, which influences the importance of precision and repeated observations in given locations.

The role of place attachment was discussed from the standpoint of place dependence in

relation to the constraints and opportunities on project design related to place. The place identity aspect of place attachment focused on the participation experience and the ways that it can influence perceptions of place among participants. Place attachment also related to concepts of ownership and sociality, which was linked to prior work on mobile computing. Ground truthing is a related data triangulation practice that reflects the relationship of place to data quality and establishing truth in science. The discussion of ground truthing revealed that it is not expertise alone that is problematic for scientists who question the quality and value of citizen science data. The inherent virtuality of participation, which makes it harder to verify that protocols are being followed correctly, compounded concerns around data quality.

The role of technologies in citizen science was considered from the perspectives of affordances and constraints. The affordances that make ICT particularly appealing to citizen science project organizers are the ability to increase the scale and scope of the project, based largely upon the use of technology-mediated communication. The cases used ICT in several ways, some of which provided unexpected benefits to projects and extended their broader impacts. ICT could also provide rewarding participation experiences, but low-cost technologies were observed to come with additional hidden costs. These constraints were primarily related to usability, data entry, and development costs.

Customization is frequently needed when projects adopt open source software or other low-cost solutions, and always comes at some expense. The discussion brought into question the assumption that technologies will attract new audiences, and highlighted the importance of usability given that older adults are frequently the majority of contributors to citizen science, as with most voluntary work. ICT has the additional potential of offering scalable ways to give feedback for further participation, which could incentivize further participation and provide reassurance to contributors who have less confidence in their performance.

Among many themes related to organizing that were observed in the cases, the discussion here focused on the organizer-contributor ratio and the relationships between institutions and sustainability. In particular, the analysis noted that the primary constraint on projects was consistently organizer time and attention, and there appears to be a practical upper limit on the number of contributors that can be supported by any one individual. Managing meta-contributors to extend the project capacity was a useful tactic for eBird, but required additional effort to put into place and worked best when a set range of structured tasks could be assigned to the individuals who could support other contributors. Additional ways that organizers attempted to address these limitations included simplifying participation expectations or increasing skill expectations to reduce demands on organizers for training and support.

These issues pointed to the importance of sustainability planning, which was often related to institutional support. Organizational settings can help sustain a project through initial development and working with partners extended the ability of project leaders to expand participation both in numbers and in geographic range. A diverse portfolio of revenue streams was also identified as a valuable strategy for supporting long-term project operation, but challenges with acquiring funding were also noted. In general, early planning for long-term sustainability was considered important for developing new projects, particularly including contingency plans for unexpected rates of growth. Further documentation of sustainability practices will be needed as the field continues to grow and develop.

Finally, links between each of these topics and concepts from both the research questions and theoretical framework were highlighted. These interconnections demonstrate the utility of the theoretical framework for focusing the inquiry to answer the research questions.

CHAPTER IX

Conclusions

This chapter begins by discussing the limitations of the study. It then reviews how the foregoing chapters answer the research questions and highlight additional relationships between concepts from the theoretical framework, drawing directly on the discussion of the emergent themes from Chapter VIII. It also suggests opportunities for future research and outlines the contributions of the current work.

9.1 Limitations

The primary limitations of the study relate to its breadth; instead, emphasis was placed on depth in order to develop a richer understanding of citizen science. These limitations are primarily related to the focus, sampling, and methods used in this work. They also suggest future directions for research, which was one of the goals of the study.

The result of the focus on organizers means that the theoretical framework is relatively one-sided. Additional concepts would undoubtedly be relevant to developing a more complete model of the phenomenon, particularly in the category of states. Including participant perspectives would also enrich the discussion of existing aspects of the framework, especially the concepts of mission, design choices, and individual development, as well as the full set of individual inputs. It would provide a deeper understanding of the alignment of scientific and personal interests, as well as the relationship of participation processes to other concepts in

the framework. Bringing in participant perspectives might also help to identify other key processes that may be absent in the theoretical framework.

The current study attempted to ameliorate this shortcoming through extensive participation and observation, but investigating the experiences of other participants and non-participants would refute, support, and extend the observations stemming from participant observation. Future work is needed to integrate the findings presented here with studies of participants. Fortunately, such research does exist, and adding complementary data from participants in the cases from this study is a feasible extension of the current work.

Decisions related to case selection posed another set of limitations to the work. The number of cases, although limited, provided opportunity for substantially greater depth than could be achieved with a larger sample. The tradeoff was made to support empirically-grounded theory development through more detailed study. Future work featuring a broader range of projects for comparison would supplement these findings, and could help establish the applicability of the findings to the larger population of citizen science projects.

In particular, the focus on observation-based citizen science projects in ecological sciences limits the transferability of the findings to other contexts. While the theoretical conceptualization specifically attempted to maintain theoretical generalizability beyond these constraints, additional research in a broader set of research domains is needed. Comparisons to citizen science projects focused on data processing tasks and projects that have no place-based elements are also needed to improve the generalizability of the framework. These are areas for future research that could be addressed with a small number of additional case studies, as some projects combine all of these elements. For example, the Zooniverse projects focus on different research domains (primarily astronomy), feature data processing tasks, and are entirely virtual in the mode of participation. Additional comparison to more diverse projects is an obvious next step for developing the research and the theoretical framework.

The limitations imposed by interviewee sampling are primarily related to the focus on organizers, as no participants were explicitly included in the interview sampling. Implicitly, however, most of the interviewees were also participants as they participated in their own citizen science projects, but brought a very different perspective on participation than would be expected of volunteers. The number of organizers who were not interviewed was minimal, with only one or two individuals missing for each case. Interviews with organizers of partnering projects and organizations helped provide additional institutional context, however, which substantially improved the diversity of interviewee sampling.

Finally, the choice of interpretive qualitative methods limits the generalizability of the study. The goal of this work was not to test hypotheses, but rather to produce findings with theoretical generalizability. The theoretical framework could offer a foundation for future confirmatory studies that draw upon more representative samples of the population of citizen science projects, although the current rate of growth in citizen science will make representative sampling a moving target for some time to come.

The interpretive approach, while well suited to the goal of deeply contextualized theory development, meant that there was no additional analytical verification, e.g., a second coder evaluating the texts to improve reliability. Ensuring dependability was instead supported throughout the research process with memos describing contextual changes and how they affected the research; these effects were minimal and tracked through ongoing contact with the case study sites. Confirmability and transferability (external validity) of the findings was supported primarily through detailed documentation of the sampling, along with thorough descriptions of the cases and researcher positionality. Negative instances were highlighted in the comparison of the cases and an audit trail was also maintained throughout the study. Several strategies were applied to strengthen the internal validity, including several rounds of peer review and multiple stages of participant review. All indications from these sources

supported the findings, verifying the credibility of the work. These findings also met the goal of addressing the research questions, discussed next.

9.2 How do virtuality and technologies alter organizing in citizen science?

The first research question focused on the ways that virtuality and technologies influence organizing processes. The case studies showed that virtuality was an inherent quality of citizen science project design. If the scientific interests of organizers could be achieved without the assistance of the broader public, they probably would not choose citizen science as a research approach due to the substantial complications introduced by involving volunteers.

As an initial condition, virtuality is part of the project processes more than it is an input like resources, which is why it does not appear in the theoretical framework as a separate concept. The facets of virtuality that were evident were spatial and temporal discontinuities of participation, which led to different organizing approaches than would have been employed in colocated research. Virtuality is in fact one of the key benefits of citizen science. The spatial and temporal spread of contributors enables new types and larger scales of scientific research. This benefit comes with challenges for organizing potentially large numbers of unknown individuals at a distance.

Notably, the ratio of organizers to contributors meaningfully influenced project design choices. Unlike large-scale scientific collaboration in which all contributors are professionals and play varying roles based on different areas of expertise, citizen science projects are typically designed so that nearly all contributors perform the same task regardless of expertise. The uniform tasks make coordination of large groups of volunteers more tractable for a small number of organizers, and help ensure better scientific rigor by requiring participants to adhere to a common data collection process.

Because the execution of these tasks cannot usually be directly observed to evaluate

performance, the quality of data that are produced by citizen science is often questioned. Mountain Watch was an exception because organizers were able to simulate and observe participation by non-experts, but all of the projects in this study employed alternate mechanisms to improve data quality. In addition, the contrast between the cases showed that the relative extent of virtuality (e.g., global, continental, or regional geographic scope) is related to the degree to which ICT is central to coordinating participation.

Digital technologies are the tool that has made large-scale virtual participation in citizen science possible. ICT reduces the coordination costs but requires creating and maintaining systems, which can be a considerable undertaking. With the right technologies, however, the number of volunteers an organizer can manage effectively increases substantially. ICT enables organizers to provide automated feedback and encouragement to contributors, increasing participant satisfaction and retention.

Technologies also influence organizing activities related to communication in other ways; the organizers of the GSP and eBird handled large volumes of email communication with participants. Some of the additional technology-mediated communication venues such as email listservs and online forums permitted knowledge sharing and problem resolution among participants, but maintaining these resources also required ongoing organizer attention.

Another aspect of technologies and organizing was the interaction of material and digital technologies with place. The requirements of the physical environment and affordances of the different tools for recording data led to the use of paper in the field, usually with subsequent online data entry. For the GSP, accepting data sheets by postal mail increased the data management burden on organizers but lowered barriers to participation, which corresponded with aspects of the project mission. The research design requirements of field-based observation posed several constraints, but also provided advantages, that would not be expected to occur in entirely technology-mediated projects without place-based elements.

Comparing the projects from this study to entirely virtual citizen science projects is an avenue for future research that would provide further insight into the tradeoffs involved in organizing projects for which participation involves different relationships to place.

As the cases also demonstrated, ICT can support data quality in several ways, including triangulation and verification of volunteer data. The plantcam images made at Mountain Watch monitoring sites can be used to evaluate volunteers' data submissions and thereby verify the reliability of data submitted by the same participants for additional locations. The eBird data review system uses a system that automatically filters submissions for review by local experts, based on meta-contributions of regional checklists, which makes data validation more scalable. eBird and the GSP both demonstrated ways in which the design of data entry interfaces can further support data quality.

9.3 How do virtuality and technologies shape participation in citizen science?

The research questions focused on the effects of virtuality and technologies on participation processes as well as organizing. Virtuality in citizen science means that participation is open to a larger and more diverse potential participant population. It also means that participants are unlikely to receive direct or extensive training on how to complete participation tasks. Relatively minimal training indicates that the tasks need to stand alone, which is usually accomplished through simplification.

With the exception of projects that are intentionally limited by habitat types or geographic ranges, most citizen science project participants are geographically distributed. Physical distribution of observers is one of the assets of this form of scientific collaboration, but also means that geographic biases are inherent. Like most people, the majority of project participants usually live in cities, and when given the choice of monitoring locations they tend to submit observations made in or near cities. In addition, some places are very thinly

populated in general, and it can be difficult to recruit participants to make observations for these areas. These geographic biases can cause problems for some research, but analytic techniques can help limit these effects.

Virtuality also means that place is an important element of these projects in several ways, particularly in terms of the functional and emotional aspects of place attachment. Functional constraints imposed by physical locations is a factor that organizers typically address through project design. Emotional relationships to places can carry both positive and negative connotations for participation. For the projects in the study, the positive aspects of subjective participant experiences of place seemed to outweigh the potential negative associations. When participants were provided autonomy to select their own places to participate, their choices contributed to geographic bias but also to convenience and satisfaction, both of which can support ongoing participation.

As previously discussed, ICT has both affordances and constraints for participation. While it makes participation possible for a larger number of people due to improved access to project resources, ICT also creates challenges for some individuals. Problems with using technologies were primarily related to usability, and potentially also to participant demographics, as project organizers reported that older adults encountered more difficulties with online data entry and account management. Despite these issues, ICT can be rewarding for participation as well. The eBird project provided many examples of ways that access to data reports and visualizations supported stronger participation and led to greater participant satisfaction and commitment.

The combination of virtuality and technologies changes the way participation is designed, and couples with both organizing and participation. Many of the findings related to organizing are also linked to participation as these processes are so extensively interrelated. Unlike most other online communities, the tasks that contributors perform are typically uniform

and assigned (rather than self-selected and unique) and coordination effects are usually reduced to pooled interdependence. Lower interdependency in task structures is part of why fewer organizers can coordinate the efforts of more contributors, and bears a stronger resemblance to crowdsourcing than other forms of distributed collaboration, particularly in science-related contexts. Future research could compare the findings from this study associated with job design and participation in citizen science with other online communities, peer production environments, and collective intelligence or crowdsourcing initiatives. Such comparison could further investigate the role of task structure and autonomy in distributed voluntary participation and the reward structures that support ongoing participation, such as system-based feedback and direct communication with organizers or other participants.

9.4 How do organizing and participation influence scientific outcomes in citizen science?

The primary purpose for citizen science is producing scientific knowledge, so the research questions examined the role of organizing and participation on scientific outcomes. The research found that these processes and scientific outcomes are inextricably linked to one another, and to science and design processes.

The case studies showed that participation had a relatively simple and direct relationship to scientific outcomes: participation is necessary to produce the contributions that are used to generate project outcomes, and greater quality and quantity of participation improves those outcomes. Scientific knowledge outcomes were evident in multiple forms, including scholarly publications as well as applied outcomes such as decision support.

Acknowledging the value of diverse types of scientific outputs is particularly important because providing a scientific basis for other activities (e.g., management or policy decisions) is often the primary goal motivating project organizers in nonacademic contexts. Future research investigating the types of science-related products of citizen science and their uses

would substantially improve the ability to evaluate project effectiveness from a scientific and organizational standpoint, and would complement existing research focused on evaluating outcomes for participants. Generating more holistic evidence of the scientific value of citizen science at the project level may also provide further support for funding based on the scientific merits of citizen science rather than just the outcomes related to individual development. While valuable, the education and outreach benefits of engaging the public are less compelling to some funders, leading to challenges in obtaining support for innovative research through citizen science projects that are less focused on informal science education.

The case studies further showed that the interaction of participation and organizing can also broaden the scope of scientific outcomes of the project. One of the causes of project evolution appears to be the participants themselves. The participants were the main resource for participation processes that improved project outcomes, and the inputs they represent are directly related to organizing and design processes, figuring into expectations of participant skills and interests. Through organizing and design processes, connections to scientific interests and science processes became apparent when participants influenced the direction and methods for the research, e.g., when the GSP contributors subjectively classified their locations leading to the incorporation of the housing density data source for analysis. Design decisions and organizing throughout initial development and ongoing operation of a citizen science project influence who participates, how committed they are, how satisfied they are, how long they continue to contribute, and whether they recruit additional participants from their own personal networks, among other participation-related effects. Each of these factors has an influence on the scientific outcomes, through the quality or quantity of observations that are contributed, or both.

As mentioned above, organizing influences the design and science processes in numerous ways that directly impact all project products, including scientific knowledge. Organizers in

the case study projects routinely evaluated data quality, sometimes using rigorous methods such as those seen in the Mountain Watch project, and subsequently made changes intended to improve the resulting data contributions and scientific products. Several approaches to improving scientific outcomes that were observed in the case studies included efforts to increase levels of participation, protocol changes to simplify or clarify procedures, varied communication strategies, broadening and narrowing of participation options, and coordination of meta-contributors to improve scalability. Resource constraints, particularly limitations on organizer time and attention, are the primary reasons that citizen science projects may experience delays or failure to produce scientific outcomes, which has important implications for project planning and sustainability.

9.5 Emergent Findings

Several other findings emerged that were not directly related to the research questions, such as the impact of design choices on the other processes in the framework. Another set of findings highlighted ways that inputs influenced all of the project processes and states. Additional factors besides virtuality and technologies had meaningful effects on products of the cases studied in this research. In particular, the emergent findings include:

1. The triadic relationship between institutions, resources, and sustainability.
2. The impacts of alignment of scientific and personal interests.
3. The implicit links between communities, organizing, and individual inputs.
4. The strategy of engaging non-scientist community members as organizers.

The triadic relationship between institutions, resources, and sustainability was evident throughout the cases. These relationships are not particularly surprising from the standpoint of organizational sociology and the connections are logical, but may not be immediately apparent to first-time citizen science project organizers. Consciously considering the

relationship of institutional and organizational settings to resources and sustainability could be a good strategy for organizers to plan more effectively for success in both the short and long term. Further research into the influence of these relationships on project goals and successes, or lack thereof, would likely have substantial benefit for project organizers in the form of new best practices and could also contribute to theory in organizational studies. This raises the issue of the value of studying failed projects, a useful direction for future work, although failures are often more difficult to identify than successes.

Viewed retrospectively, the impacts of the alignment of scientific and personal interests also seems logical. The importance of this element of project design, while often implicit in practice, was not self-evident until the Deliberate Design Model was introduced into the theoretical framework, as it was also clearly supported by the empirical data. While a strong correspondence between these qualities is desirable for all stakeholders in a citizen science project, it may not be possible to achieve in every situation. When it becomes apparent that the alignment of scientific and personal interests will be partial at best, project organizers have the simultaneous challenge and opportunity of finding other means to support participation and scientific outcomes. Research into the relationship of goals and motivations for both participants and organizers could lead to new theoretical insights into the more detailed interactions between these concepts and, for example, communication strategies. Determining ways to evaluate the match of participant and organizer interests would be a useful contribution to practice.

The links between communities, organizing, and individual inputs are another area of implicit understanding among project organizers that was made explicit through the analysis presented in this study. While the relationships between community and individual inputs were not strongly stressed by interviewees, the lack of emphasis on this connection could be due to the taken-for-granted nature of self-selection of individuals into communities and the

related requirements for community membership. The influence of the degree to which organizers understand both the relevant communities and personal characteristics of participants was also apparent through the comparison of eBird to the other two projects. The eBird project leaders' intimate engagement with the birding community meant that they were able to effectively manage project and system development initiatives.

The strategy of engaging non-scientist community members as organizers was a clear effort to counter the challenges that arise when there is a less complete understanding of the contributor base on the part of the organizers. Community-centric staffing was also successful in improving outcomes to the degree that organizers were able to act on new insights from working closely with individuals who are deeply embedded in the communities from which participants are recruited. Investigating whether the strategy has been employed more broadly in citizen science, and the impacts it has on project processes and outcomes, would verify whether this a best practice for project management. Better understanding the specific qualities of community representatives that make them effective organizers would also have potential for practical and theoretical value, and would be interesting to compare to other types of contributory projects.

9.6 Contributions

This study made both theoretical and practical contributions. The empirically-based theoretical framework both complements and extends prior models, such as the Deliberate Design Model. It provides a foundation for future theoretically motivated research. More specifically, the theoretical framework has room for expansion and refinement through incorporation of participant perspectives and case studies of citizen science in different domains, in place-independent contexts, or in which participation is designed around data processing tasks.

The framework has potential for application to other forms of scientific collaboration because it contains three concepts specific to a scientific context, scientific interests as inputs, science processes, and scientific knowledge outcomes. There is also potential to apply this framework to other forms of contributory communities, by substituting other contextually-relevant inputs, processes, and outputs for those science-focused concepts. Interestingly, none of the concepts that emerged as important in this study are specific to distributed collaboration or online communities. This is partly because technologies is a broader concept than just information and communication technologies, so it also refers to tools like scientific protocols and material technologies like paper data sheets.

The practical contributions of the work include the identification of new prospective best practices. It surfaces taken-for-granted relationships between important elements of citizen science project design and management, and provides examples of the impacts of related choices and strategies on project outcomes. The theoretical framework could also support expansion and further development of current heuristics for project planning and evaluation.

In addition, the comparative case study offers in-depth descriptions and comparison of citizen science projects that engage the public in observation and data collection tasks. The descriptions and analysis can provide useful points of reference for practitioners. The case studies also form a basis for comparison of citizen science to other types of technology-supported contributory communities. The focus on organizers complements other research centered on participant viewpoints, and opens up opportunities for merging these perspectives to foster a more comprehensive understanding of the phenomenon.

9.7 Summary

Citizen science is rapidly increasing in popularity as a method for achieving scientific goals that were previously out of reach. Citizen science projects offer potential for innovative

scientific research, broader impacts, and personal development for participants. The success of these projects hinges on understanding the implications of the relationships between initial conditions and resources, ongoing processes, and the resulting impacts on outcomes.

Through the development of an empirically-grounded theoretical framework, rich descriptions of three citizen science projects, and comparative analysis of these cases, this research has elucidated several important aspects of project design and management. The study reflects new opportunities for better understanding both citizen science and technology-supported contributory communities, illustrates the value of a theoretical framework for shedding light on aspects of citizen science projects that have been taken for granted, and presents a multidisciplinary view on a phenomenon that has not previously been considered from this perspective. Foundational studies like the one presented here can help to support future research on citizen science as well as best practices in the community of citizen science organizers.

APPENDICES

APPENDIX A

Sample Interview Protocol

1. Could you tell me about your professional background and how you became involved with the project?
2. How is your work connected with the project?
3. How have you been involved with the project's...?
 - (a) project design and adaptation
 - (b) data management and validation
 - (c) volunteer and community management
4. How do technologies fit into project participation?
5. Is there anything else that you think is important about the project that we haven't discussed?

APPENDIX B

Survey Methods

A survey instrument was composed to directly elicit selected descriptive characteristics of projects; this section describes the design of the survey instrument, sampling frame, and response rate.

The survey instrument was presented as a two-part questionnaire: first, a brief project profile and second, a separate, lengthier survey. The first portion of the questionnaire was a project profile, allowing projects to opt-in for listing on several cooperating websites that provide listings of citizen science projects, and update existing project profiles based on data provided with the sampling frame or create a new project profile (Appendix C). The second portion of the questionnaire was the project survey, which asked for additional details in several categories. The full survey included 57 items, including both multiple choice questions and free-response spaces for each structured item (Appendix D). The answers for multiple choice items were developed from existing public data on citizen science projects and a prior study that provided a foundation for the survey by identifying over 80 characteristics of citizen science projects based on a purposive sample of 30 projects (Wiggins & Crowston, 2010). There were no required fields, so each item had a variable response rate. The items covered several categories, but those reported in this paper focused on data validation methods.

The sampling frame was composed of projects listed on Cornell Lab of Ornithology's citizen science email list and in the now-defunct Canadian Citizen Science Network. These are the most comprehensive sources of contacts for North American citizen science projects. Approximately 60 additional contacts were manually mined from the online community directory at <http://www.scienceforcitizens.net> to extend the disciplinary diversity of the sample.

These sources provided a combined set of approximately 840 contacts after removing duplicates and bad addresses. These contacts are individuals who had self-identified as responsible for or involved in the management of citizen science projects. Approximately 280 projects were identified in this process, and another 560 individuals who may be connected with additional projects were also invited to participate.

In response to approximately 840 emailed requests for participation, 128 project profiles were created or updated. Seventy-three surveys were initiated and 63 fully completed, for a participation rate of 15% and a response rate of approximately 8%. The surveys and profiles were combined for analysis. The response rate is low, though not atypical for such a survey. However, it should be noted that the number of projects is smaller than the number of contacts, meaning that the response rate for projects (our unit of analysis) is better than it appears. As noted above, we were able to identify approximately 280 projects, which would lead to a response rate of about 22% rather than 8%; the actual response rate lies somewhere in between these two figures.

Most of the responses came from small-to-medium sized projects, based in the United States, with several Canadian projects reporting, and three from the UK; a handful of projects are organized by research teams that span international boundaries. Nearly all responding projects are of the monitoring and observation types. The sample is also subject to self-selection bias, such that projects interested in attracting more participants through a directory listing were more likely to respond than those that may selectively engage contributors, for example, based on known subject expertise. However, despite these limitations, it is believed that the resulting sample is generally representative of the population of citizen science projects.

APPENDIX C

Citizen Science Project Profile

Opt-in

Let us know whether to include your project's profile in our partner sites' citizen science project directories whenever appropriate by selecting them above. At least one selection is required.

If you select "None", your project's profile will not appear in any of our partners' project directories, and any existing project listings will be removed. Your response will still be valuable for helping us to better understand the characteristics and needs of citizen science projects.

- Citizen Science Central
- Science for Citizens
- USA National Phenology Network
- DataONE
- Additional partners that may be added in the future
- None

Project contact information

Project name

Project website

Please enter the primary website for the project, starting with <http://>. If the project does not have a website, enter "None".

Contact name

Name of a contact person for the project.

Contact email

Please provide a general contact email address. This does not need to be the same email address that was used to create your account.

Project partners

Organizations or groups that partner to organize and manage this project; please separate each group name with a comma.

Affiliated websites

Please enter any other websites related to this project, such as web sites for organizing groups, starting with http://.

RSS feed

If your project provides an RSS feed for news or updates, include it here to have your updated content featured by our partners.

Project logo

If you would like to include a logo or image to represent your project, please upload it here. You can upload a file up to 1 MB, for any of these file types: gif, jpg, png, bmp, tif, and pdf.

Project Description**Year project started**

What year did the project start? Format: 2010

Project description

Briefly describe your project (about a paragraph).

Project subject categories

Which subject categories are related to your project's focus? Please check all the general subject areas that apply.

- Animals
- Archeology
- Astronomy & Space
- Birds

- Biology
- Chemistry
- Climate & Weather
- Computers & Technology
- Ecology & Environment
- Food
- Geology & Earth Science
- Health & Medicine
- Insects
- Nature & Outdoors
- Ocean & Water
- Physics
- Science Policy
- Sound
- None of the above

Project topic keywords

Please enter topic keywords that describe your project more specifically; separate keywords with commas.

Examples: Invasive species, worms, light pollution, water quality, weather, phenology, plants, possums, climate change, marine, etc.

Project Participation

Audiences

What audiences do you try to engage? Please check all that apply.

- Students
- Families
- Youth
- Adults
- Teachers
- Landowners
- Retirees
- Enthusiasts

- Community groups
- Schools
- Nature centers
- All of the above

Other audiences

Please list any additional audience groups for your project.

Geographic scope

What is the approximate geographic range of participation?

Examples: Global, international, North America, northeastern US, provincial, New York state, tri-county area, city, etc.

Indoors or Outdoors

Does participation involve activities that are done indoors or outdoors? If both indoors and outdoors locations may be involved, please select both items.

- Indoors
- Outdoors

Place

Where do project activities occur for most participants?

- Online exclusively
- Anywhere participants choose
- Anywhere within a specific range
- At one or more specific locations
- Other

Project duration

How is the timing of the project structured? Is it more like an annual event or campaign (e.g., Christmas Bird Count), a one-time effort that doesn't repeat year after year (e.g., a one-time bioblitz), or an ongoing project with either year-round or seasonal participation?

- Annual event
- One-time limited duration project
- Ongoing project, year-round
- Ongoing project, seasonal
- Other

Learning materials

Does the project provide learning materials? Please check all that apply.

- None
- Lesson plans
- Other classroom teaching materials
- Family learning activities
- Training materials
- Links to other resources

Training requirements

Please briefly describe any training requirements.

Examples: None, online quiz, self-paced tutorial, short training session (up to 3 hours), workshop series (three 4-hour workshops), etc.

Required gear

Please list any *required* equipment or tools that contributors need in order to participate effectively; separate each item with a comma.

Examples: binoculars, smartphone, maps, water monitoring kit, global positioning device, sweep net, computer, etc.

APPENDIX D

Citizen Science Project Survey

This survey is being conducted by researchers at Syracuse University to better understand the characteristics and needs of citizen science projects. Many of the questions will help determine the direction of current efforts to better support citizen science projects.

Individual responses to these questions will not be visible to anyone but you and our research team. Aggregated responses will be shared with interested participants and collaborators.

There are no required responses, and you can stop the survey at any time. You can also save your survey answers and return to complete the survey, but please don't delay - this is a one-time survey, and will only be available through the middle of March 2011.

Project name

Please re-enter the project name so that we can associate the project profile information with the additional detail from your survey responses.

Is this project currently active?

- Yes
- No
- Don't know/not sure

Project Resources

The questions in this section will help us understand the resources available to citizen science projects, which will help develop recommendations for tools and services to prioritize for future development.

Project Staffing

Please indicate the number of full-time equivalent (FTE) paid employees organizing or running this project. If there are no paid staff for the project, please enter “0”.

Example: 2.5 FTE

Annual Operating Budget

What is the approximate annual operating budget for the project? Please note currency type if not in US Dollars, e.g. CAD, GBP, EUR, etc.

Funding Sources

What sources of funding support the project? Please check all that apply.

- Participant fees
- Federal grants
- Other grants
- Private donations
- Sponsorships
- Licensing
- Service fees
- Memberships
- Merchandise sales
- Advertising
- In-kind contributions
- Not sure/don't know

Other Funding Sources

Please list any additional or more specific funding sources for your project.

Comments on Project Resources

Please feel free to include any additional comments about project resources.

Participation Details

The questions in this section will help us better understand the types of activities and forms of participant engagement that projects need to be able to effectively support.

Participation Activity Types

What are the primary types of activities for people contributing to the project? Please check all that apply.

- Observation
- Species identification
- Classification or tagging
- Data entry
- Finding entities (e.g., in images, in natural habitats)
- Measurement
- Specimen/sample collection
- Sample analysis
- Site selection and/or description
- Geolocation
- Photography
- Data analysis

Other Participation Activities

Please list any additional types of activities that contributors participate in.

Rewards to Contributors

Are there any explicit material or status rewards for participants? Please check all that apply.

- None
- Free equipment/supplies/training
- Certificate
- T-shirts
- Promotional items, e.g. stickers, pins, keychains, patches
- Top contributor listings
- Personal performance ratings
- Public acknowledgment
- Role advancement
- Editor/moderator privileges
- Naming privileges
- Co-authorship privileges
- Volunteer appreciation events

Other Rewards

Please list any additional types of explicit rewards that contributors may receive from participating.

Social Opportunities

What opportunities for social interaction are available to participants? Please check all that apply.

- None
- Forums
- Email listservs
- Blogging and/or commenting on blogs
- Social media (e.g., Twitter, Facebook, etc.)
- Conference calls or webinars
- Meetings
- Training sessions
- Volunteer appreciation events
- Group participation in project activities
- Classroom participation

Other Social Opportunities

Please list any additional opportunities for socializing among participants.

Comments on Participation Details

Please feel free to share any additional comments about project participation activities, rewards, and social opportunities, or other topics related to participation.

Tools and Technologies

The items in this section will help us better understand the current state and future plans for tools and technologies to support citizen science projects, which will be used for recommendations on infrastructure planning and resource development.

Communication Tools

What types of communication tools and technologies does the project use? Please select all that apply.

- None
- Website
- RSS
- Email
- Conference calls or webinars
- Print publications
- Research articles
- Blogs
- Forums
- Photo galleries
- Maps
- Graphs and charts
- Animated or interactive data visualizations
- Data querying and summary tools
- Social media (e.g., Twitter, Facebook)

Other Communication Tools

Please list any additional communication tools or technologies the project uses.

Technology Plans

Thinking more broadly than just tools for communication, what new technologies does your project plan to implement in the next two years? Please briefly describe any plans for technology changes or additions.

Future Technologies

What new technologies or improvements to your current technologies would you like to implement in the future, beyond what is currently planned?

Comments on Tools and Technologies

Please feel free to include any additional comments about project tools and technologies.

Data Management

These questions will help us better understand the data management and policy needs of citizen science projects, allowing us to identify priorities for future infrastructure investment.

Data Validation Methods

What methods of validation or quality control are used? Please check all that apply.

- None
- Expert review
- Automatic filtering of unusual reports
- Replication or rating, by multiple participants
- Replication or rating, by the same participant
- Photo submissions
- Paper data sheets submitted along with online entry
- Rating of established control items
- Uniform equipment
- QA/QC training program
- Validation planned but not yet implemented
- Not sure/don't know

Other Validation Methods

Please describe any additional validation methods used in your project.

Data Sharing

With whom does the project currently share data? Please check all that apply.

- No data sharing
- Sharing with contributors
- Sharing with project-affiliated researchers
- Sharing with a research network or data archive
- Sharing with the general public
- Sharing is planned but not yet in place
- Not sure/don't know

Other Data Sharing

Please list any additional specific groups with whom the project shares data.

Data Ownership

What is the project's policy on data ownership? Choose the options that best fit, or describe a different arrangement below.

- No policy
- Currently developing policy
- Researchers own the data
- Project contributors own the data
- Third party owns the data
- Public owns the data
- Not sure/don't know

Other Data Owners

Please describe the project's data ownership policy if it is different from the types above.

Comments on Data Management

Please feel free to share any additional comments on data management and policies.

Project Contributions

Because there are many different ways to describe contributions and contributors, we ask you to define the primary unit of contribution to your project (and list secondary types of contributions) below. The definition of a contributor in the following questions is a person who has made any such contribution to the project.

These questions will help us better understand the diversity of citizen science projects with respect to types and rates of contribution. If you don't have details handy, please give your best approximation.

Unit of Contribution

What is the unit of contribution for this project?

Examples: observations, specimens, samples, classifications, images

Other Contributions

Please list any additional or alternate forms of contribution to the project.

Examples: blog posts, forum or blog comments, additional protocols, mentoring, etc.

2010 Registrations

Approximately how many people have registered or signed up for the project in 2010?

Total Registrations to Date

Approximately how many people have registered or signed up for the project to date?

2010 Contributors

How many people made a contribution to the project in 2010 by submitting data, completing tasks, or other active engagement?

Total Contributors to Date

Of those who have registered or signed up for the project, how many have made a contribution to the project by submitting data, completing tasks, or other active engagement?

2010 Contributions

Approximately how many contributions have been made to the project in 2010?

Total Contributions to Date

Approximately how many contributions have been made to the project to date?

Comments on Contributions

Please feel free to share any comments about project contributions, contributors, and meaningful measures of participation.

Project Goals & Outcomes

This set of questions will help us better understand the range of goals and desired outcomes that are important to citizen science projects.

Project Goals

Please indicate how important each type of goal is for your project. Any other goals you consider important that are missing here can be indicated below.

Answer options: Not important, Low importance, Slightly important, Neutral, Moderately important, Very important

Goal types, presented in random order:

- Science
- Management
- Action
- Education

- Conservation
- Monitoring
- Restoration
- Outreach
- Stewardship
- Discovery

Other project goals

Include any other project goals here, and please feel free to clarify the relative importance of project goals.

Intended Project Outcomes

What outcomes does the project intend to produce? Please select all that apply.

- Data sets
- Data analysis
- Academic publications and presentation
- Technical reports
- New discoveries
- New research methods
- New inquiry
- Policy changes
- Community action
- Environmental restoration
- Individual learning

Other Intended Outcomes

Please list any additional intended project outcomes that are not included above.

Actual Project Outcomes

What are the actual project outcomes to date? Please select all that apply.

- Data sets
- Data analysis
- Academic publications and presentation
- Technical reports

- New discoveries
- New research methods
- New inquiry
- Policy changes
- Community action
- Environmental restoration
- Individual learning

Other Actual Outcomes

Please list any additional project outcomes to date that are not included above.

Project Evaluations

Have any of these types of evaluations been conducted for your project? This could include any efforts to gather data about project participants and their needs, information about whether the project is working well, or evidence about project impacts.

Answer options:

- Yes, definitely
- Yes, I think so
- No, I don't think so
- No, definitely not
- Don't know

Evaluation types:

- Front-end (needs assessment or baseline information)
- Formative or process (during project development or implementation)
- Summative (describing project outcomes or impacts)

Comments on Project Outcomes

Please feel free to share any additional comments on project evaluation and project outcomes.

Participant Outcomes

These questions will help us understand the participant outcomes that projects hope to achieve, and will be used to prioritize the development of assessment tools for better understanding participant outcomes.

Science Knowledge Outcomes

Of the outcomes listed here, which is the most important to your project for assessing how participants' *science knowledge* changed as a result of participation?

- Increased knowledge of specific science content
- Increased knowledge of the process of science (i.e., methods used in science)
- Increased knowledge of the nature of science and the scientific enterprise
- Other (please describe below)

Other Science Knowledge Outcomes

If you marked “Other” in the previous question, please use this space to describe the outcome you would like to assess.

Science Interest Outcomes

Of the outcomes listed here, which is the most important to your project for assessing how participants' *interest in science* changed as a result of participation?

- Increased interest in specific science issues
- Increased interest in science careers
- Increased interest in nature/environment
- Other (please describe below)

Other Science Interest Outcomes

If you marked “Other” in the previous question, please use this space to describe the outcome you would like to assess.

Science Skills Outcomes

Of the outcomes listed here, which is the most important to your project for assessing how participants' *scientific skills* changed as a result of participation?

- Improved ability to ask scientific questions
- Improved ability to identify, collect, and submit accurate data
- Improved ability to analyze and interpret data
- Improved use of technology
- Other (please describe below)

Other Science Skills Outcomes

If you marked “Other” in the previous question, please use this space to describe the outcome you would like to assess.

Attitude Outcomes

Of the outcomes listed here, which is the most important to your project for assessing how participants' *attitudes* changed as a result of participation?

- Improved attitudes toward science
- Improved attitudes toward self as a scientist
- Improved attitudes toward nature/environment
- Other (please describe below)

Other Attitude Outcomes

If you marked "Other" in the previous question, please use this space to describe the outcome you would like to assess.

Behavior Outcomes

Of the outcomes listed here, which is the most important to your project for assessing how participants' *behaviors* changed as a result of participation?

- Increased citizen action/involvement with policy
- Increased participation with science-based activities
- Increased environmental stewardship
- Other (please describe below)

Other Behavior Outcomes

If you marked "Other" in the previous question, please use this space to describe the outcome you would like to assess.

Conclusion

Other Considerations

Are there any thoughts you'd like to share with us about your project or this survey?

Follow-up Studies

Due to the exciting growth of citizen science, several research and infrastructure projects will be seeking additional input for documenting best practices, developing recommendations, and planning infrastructure to support citizen science projects.

Please select an option to let us know whether you are open to participating in future surveys, interviews or focus groups that will help shape the direction of policy and resource development. Indicating your interest does not create any obligation for future response, and only means that you would consider participating.

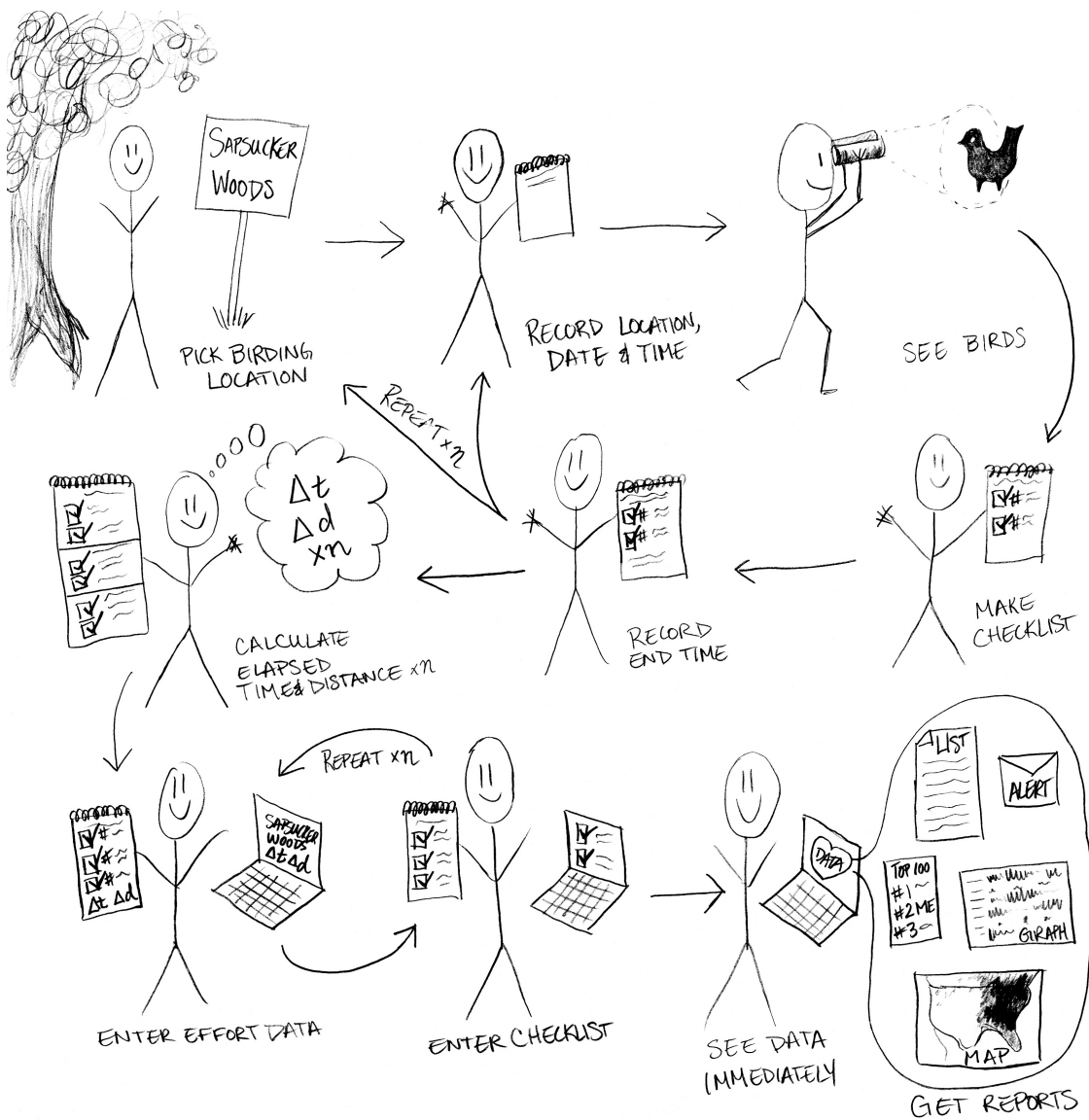
- Yes, I would consider participating in future studies.
- No, I am not interested in participating.

Survey Results

If you completed the survey and would like to receive a summary of the results, please enter an email address to which we can send it. Any email address you provide here will be used only once, when we send you the survey results summary, which should be available by June 2011.

APPENDIX E

Sample Rich Process Model



BIBLIOGRAPHY

BIBLIOGRAPHY

- Adam, D., Tarrow, S., & Tilly, C. (1997). Toward an integrated perspective on social movements and revolution. In *Comparative Politics: Rationality, Culture, and Structure* (pp. 142–173). New York, NY: Cambridge University Press.
- Allison-Bunnell, S., & Thompson, S. (2007). Débutants et experts dans la science citoyenne nord-américaine: une théorie de la pratique,. In F. Charvolin, A. Micoud, & L. Nyhard (Eds.), *Des Sciences Citoyennes? La Question de L'Amateur Dans les Sciences Naturalistes*. Sainte-Etienne, France.
- Almirall, E., & Wareham, J. (2008). Living Labs and open innovation: Roles and applicability. *The Electronic Journal for Virtual Organizations and Networks*, 10(3), 21–46.
- Argyris, C., & Schon, D. (1989). Participatory action research and action science compared: A commentary. *American Behavioral Scientist*, 32(5), 612–23.
- Bäckstrand, K. (2003). Civic science for sustainability: Reframing the role of experts, policy-makers and citizens in environmental governance. *Global Environmental Politics*, 3(4), 24–41.
- Bailey, C. (2007). *A Guide to Qualitative Field Research*. Thousand Oaks, CA: Pine Forge Press.
- Ballard, H., Pilz, D., Jones, E., & Getz, C. (2005). *Training Curriculum for Scientists and Managers: Broadening Participation in Biological Monitoring*. Corvallis, OR: Institute for Culture and Ecology.
- Ballard, H., Trettevick, J., & Collins, D. (2008). Comparing participatory ecological research in two contexts: An immigrant community and a native American community on Olympic Peninsula, Washington. In C. Wilmsen, W. Elmendorf, L. Fisher, J. Ross, B. Sararthy, & G. Wells (Eds.), *Partnerships for empowerment: Participatory research for community-based natural resource management* (pp. 187–215). London, UK: Routledge.
- Bannon, L. J., & Schmidt, K. (1989). CSCW: Four characters in search of a context. In *Proceedings of the First European Conference on Computer supported cooperative work* (pp. 358–372). Gatwick, London, UK.

- Baretto, C., Fastovsky, D., & Sheehan, P. (2003). A model for integrating the public into scientific research. *Journal of Geoscience Education*, 50(1), 71-75.
- Bauer, M., Petkova, K., & Boyadjieva, P. (2000). Public knowledge of and attitudes to science: Alternative measures that may end the “Science War”. *Science Technology & Human Values*, 25(1), 30-51.
- Baum, F., MacDougall, C., & Smith, D. (2006). Participatory action research. *Journal of Epidemiology and Community Health*, 60(10), 854.
- Benkler, Y. (2002). Coase’s penguin, or, Linux and the nature of the firm. *Yale Law Journal*, 112(3), 367–445.
- Best, S., & Krueger, B. (2004). *Internet Data Collection*. Thousand Oaks, CA: Sage Publications.
- Bhattacharjee, Y. (2005). Citizen scientists supplement work of Cornell researchers. *Science*, 308(5727), 1402-1403.
- Birnholtz, J., & Bietz, M. (2003). Data at work: Supporting sharing in science and engineering. In *Proceedings of the 2003 International ACM Conference on Supporting Group Work* (pp. 339–348). New York, NY.
- Bonney, R., Ballard, H., Jordan, R., McCallie, E., Phillips, T., Shirk, J., & Wilderman, C. (2009). *Public Participation in Scientific Research: Defining the Field and Assessing Its Potential for Informal Science Education. A CAISE Inquiry Group Report* (Tech. Rep.). Washington, DC: Center for Advancement of Informal Science Education (CAISE).
- Bonney, R., Cooper, C., Dickinson, J., Kelling, S., Phillips, T., Rosenberg, K., & Shirk, J. (2009). Citizen science: A developing tool for expanding science knowledge and scientific literacy. *BioScience*, 59(11), 977–984.
- Bonney, R., & LaBranche, M. (2004). Citizen science: Involving the public in research. *ASTC Dimensions*, 13.
- Bos, N., Zimmerman, A., Olson, J., Yew, J., Yerkie, J., Dahl, E., & Olson, G. (2007). From shared databases to communities of practice: A taxonomy of collaboratories. *Journal of Computer-Mediated Communication*, 12(2), 652–672.
- Bowker, G. (2005). *Memory Practices in the Sciences*. Cambridge, MA: MIT Press.
- Bowker, G., & Star, S. (2000). *Sorting Things Out: Classification and Its Consequences*. The MIT Press.
- Brannick, T., & Coghlan, D. (2007). In defense of being “native”: The case for insider academic research. *Organizational Research Methods*, 10(1), 59–74.

- Brossard, D., Lewenstein, B., & Bonney, R. (2005). Scientific knowledge and attitude change: The impact of a citizen science project. *International Journal of Science Education, 27*(9), 1099-1121.
- Bryant, S., Forte, A., & Bruckman, A. (2005). Becoming Wikipedian: Transformation of participation in a collaborative online encyclopedia. In *Proceedings of the 2005 International ACM Conference on Supporting Group Work* (p. 10). New York, NY.
- Buckland, M. (1991). Information as thing. *Journal of the American Society for Information Science, 42*(5), 351-360.
- Buni, C. (2012). Still making us stretch. *AMC Outdoors Magazine, 78*, 28-33.
- Butler, B. (2001). Membership size, communication activity, and sustainability: A resource-based model of online social structures. *Information Systems Research, 12*(4), 346-362.
- Chin, G., Jr., & Lansing, C. (2004). Capturing and supporting contexts for scientific data sharing via the biological sciences collaboratory. In *Proceedings of the 2004 ACM conference on Computer supported cooperative work* (pp. 409-418). New York, NY.
- Cho, A., & Clery, D. (2009). Astronomy Hits the Big Time. *Science, 323*(5912), 332.
- Chompalov, I., Genuth, J., & Shrum, W. (2002). The organization of scientific collaborations. *Research Policy, 31*(5), 749-767.
- Chudoba, K., Wynn, E., Lu, M., & Watson-Manheim, M. (2003). How virtual are we? Measuring virtuality and understanding its impact in a global organization. In *Proceedings of the International Conference on Information Systems*. Atlanta, GA.
- Clark, F., & Illman, D. (2001). Dimensions of civic science: Introductory essay. *Science communication, 23*(1), 5.
- Cnaan, R. A., & Cascio, T. A. (1999). Performance and commitment: Issues in management of volunteers in human service organizations. *Journal of Social Service Research, 24*, 1-38.
- Cohen, J. (2000). Balancing the collaboration equation. *Science, 288*(5474), 2155.
- Cohn, J. (2008). Citizen science: Can volunteers do real research? *BioScience, 58*(3), 192-107.
- Compeau, D., & Higgins, C. (1995). Computer self-efficacy: Development of a measure and initial test. *MIS Quarterly, 189*-211.

- Cooper, C., Dickinson, J., Phillips, T., & Bonney, R. (2007). Citizen science as a tool for conservation in residential ecosystems. *Ecology and Society*, 12(2).
- Cox, A., Clough, P., & Marlow, J. (2008). Flickr: A first look at user behaviour in the context of photography as serious leisure. *Information Research*, 13.
- Crawford, S., & Ostrom, E. (1995). A grammar of institutions. *American Political Science Review*, 582–600.
- Crowston, K. (2000). Process as theory in information systems research. In *Organizational and social perspectives on information technology* (p. 149). Aalborg, Denmark.
- Crowston, K., Annabi, H., Howison, J., & Masango, C. (2005). Effective work practices for FLOSS development: A model and propositions. In *Proceedings of the Hawai'i International Conference on System Sciences*. Big Island, Hawai'i.
- Crowston, K., & Fagnot, I. (2008). *The motivational arc of massive virtual collaboration*. Lüneberg, Germany.
- Crowston, K., Li, Q., Wei, K., Eseryel, U., & Howison, J. (2007). Self-organization of teams for Free/Libre Open Source software development. *Information and Software Technology*, 49(6), 564–575.
- Crowston, K., Wei, K., Li, Q., & Howison, J. (2006). Core and periphery in Free/Libre and Open Source software team communications. In *Proceedings of the Hawai'i International Conference on System Sciences*. Washington, DC.
- Cummings, J., & Cross, R. (2003). Structural properties of work groups and their consequences for performance. *Social Networks*, 25(3), 197–210.
- Cummings, J., & Kiesler, S. (2005). Collaborative research across disciplinary and organizational boundaries. *Social Studies of Science*, 35(5), 703–722.
- Cyert, R., & March, J. (1963). *A Behavioral Theory of the Firm*. Prentice-Hall, Inc.
- Daft, R., & Weick, K. (1984). Toward a model of organizations as interpretation systems. *Academy of Management Review*, 9(2), 284–295.
- Danielsen, F., Burgess, N., & Balmford, A. (2005). Monitoring matters: Examining the potential of locally-based approaches. *Biodiversity and Conservation*, 14(11), 2507–2542.
- Danielsen, F., Burgess, N., Balmford, A., Donald, P., Funder, M., Jones, J., Alviola, P., Balete, D., Blomley, T., Brashares, J., et al.. (2009). Local participation in natural resource monitoring: A characterization of approaches. *Conservation Biology*, 23(1), 31–42.

- David, P., den Besten, M., & Schroeder, R. (2006). How 'open' is e-science? In *Proceedings of the IEEE 2nd International Conference on eScience and Grid Computing*. Washington, DC.
- Davies, C. (2008). *Reflexive Ethnography: A Guide to Researching Selves and Others*. New York, NY: Routledge.
- Davis, F. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3), 319–340.
- de Solla Price, D. (1963). *Little Science, Big Science* (Vol. 292). Columbia University Press.
- Delaney, D., Sperling, C., Adams, C., & Leung, B. (2007). Marine invasive species: Validation of citizen science and implications for national monitoring networks. *Biological Invasions*, 1573–1464.
- De Roure, D., Goble, C., Bhagat, J., Cruickshank, D., Goderis, A., Michaelides, D., & Newman, D. (2008). myExperiment: Defining the social virtual research environment. 182–189.
- Dickinson, J., Zuckerberg, B., & Bonter, D. (2010). Citizen science as an ecological research tool: Challenges and benefits. *Annual Review of Ecology, Evolution, and Systematics*, 41(1), 148–172.
- Dufour, C., & Crisfield, E. (Eds.). (2008). *The Appalachian Trail MEGA-Transect*. Harpers Ferry, WV.
- Eisenhardt, K. (1989). Building theories from case study research. *Academy of Management Review*, 532–550.
- Ellis, D., Oldridge, R., & Vasconcelos, A. (2004). Community and virtual community. *Annual Review of Information Science and Technology*, 38, 145–188.
- Emerson, R., Fretz, R., & Shaw, L. (1995). *Writing Ethnographic Fieldnotes*. Chicago, IL: University of Chicago Press.
- Eriksson, M., Niitamo, V., & Kulkki, S. (2005). State-of-the-art in utilizing Living Labs approach to user-centric ICT innovation—A European approach. *Development*.
- Farooq, U., Ganoe, C., Carroll, J., & Giles, C. (2007). Supporting distributed scientific collaboration: Implications for designing the CiteSeer collaboratory. In *Proceedings of the Hawai'i International Conference on Systems Sciences* (Vol. 40, p. 428). Washington, DC.

- Fernandez-Gimenez, M., Ballard, H., & Sturtevant, V. (2008). Adaptive management and social learning in collaborative and community-based monitoring: A study of five community-based forestry organizations in the western USA. *Ecology and Society*, 13(2), 4.
- Fetterman, D. (1998). *Ethnography: Step by step*. Thousand Oaks, CA: Sage Publications, Inc.
- Finholt, T. (2002). Collaboratories. *Annual Review of Information Science and Technology*, 36, 73–107.
- Firehock, K., & West, J. (1995). A brief history of volunteer biological water monitoring using macroinvertebrates. *Journal of the North American Benthological Society*, 14(1), 197–202.
- Fischer, G. (2002). Beyond ‘couch potatoes’: From consumers to designers and active contributors. *First Monday*, 7(12-2).
- Fogg, B., Marshall, J., Laraki, O., Osipovich, A., Varma, C., Fang, N., Paul, J., Rangnekar, A., Shon, J., Swani, P., et al.. (2001). What makes Web sites credible?: A report on a large quantitative study. In *Proceedings of the International Conference on Human Factors in Computing Systems* (pp. 61–68). New York, NY.
- Galbraith, J. (1974). Organization design: An information processing view. *Interfaces*, 4(3), 28–36.
- Ganz, M. (2000). Resources and resourcefulness: Strategic capacity in the unionization of California agriculture, 1959–1966. *American Journal of Sociology*, 1003–1062.
- Giddens, A. (1979). *Central problems in social theory: Action, structure, and contradiction in social analysis*. University of California Press.
- Gieryn, T. (2002). Three truth-spots. *Journal of History of the Behavioral Sciences*, 38(2).
- Goodhue, D., & Thompson, R. (1995). Task-technology fit and individual performance. *MIS Quarterly*, 213–236.
- Gotham, K. (1999). Political opportunity, community identity, and the emergence of a local anti-expressway movement. *Social Problems*, 46, 332.
- Graham, E., Henderson, S., & Schloss, A. (2011). Using mobile phones to engage citizen scientists in research. *Eos, Transactions, American Geophysical Union*, 92(38), 313.

- Graham, J., Simpson, A., Crall, A., Jarnevich, C., Newman, G., & Stohlgren, T. (2008). Vision of a cyberinfrastructure for nonnative, invasive species management. *BioScience*, *58*(3), 263–268.
- Granovetter, M. (1973). The strength of weak ties. *American Journal of Sociology*, 1360–1380.
- Granovetter, M. (1985). Economic action and social structure: The problem of embeddedness. *American Journal of Sociology*, 481–510.
- Grudin, J. (1994). Groupware and social dynamics: Eight challenges for developers. *Communications of the ACM*, *37*(1), 92–105.
- Guzzo, R. A., & Dickson, M. W. (1996). Teams in organizations: Recent research on performance and effectiveness. *Annual Review of Psychology*, *47*, 307–338.
- Hackman, J. (1987). The design of work teams. In *Handbook of organizational behavior* (Vol. 315, p. 342). Prentice-Hall.
- Hackman, J., & Kaplan, R. (1974). Interventions into group process: An approach to improving the effectiveness of groups. *Decision Sciences*, *5*, 459–480.
- Hackman, J. R., & Morris, C. G. (1978). Group tasks, group interaction process, and group performance effectiveness: A review and proposed integration. In L. Berkowitz (Ed.), *Group Processes* (Vol. 8, pp. 45–99). New York: Academic Press.
- Harrison, S., & Dourish, P. (1996). Re-place-ing space: The roles of place and space in collaborative systems. In *Proceedings of the 1996 ACM conference on Computer supported cooperative work* (pp. 67–76). New York, NY: ACM.
- Haythornthwaite, C. (2009). Crowds and communities: Light and heavyweight models of peer production. In *Proceedings of the hawai'i international conference on systems sciences*. Washington, DC.
- Heidegger, M. (1977). *The Question Concerning Technology and Other Essays* (W. Lovitt, Trans.). New York, NY: Harper & Row.
- Hey, T., & Trefethen, A. E. (2005). Cyberinfrastructure for e-science. *Science*, *308*(5723), 817–821.
- Hidalgo, M., & Hernandez, B. (2001). Place attachment: Conceptual and empirical questions. *Journal of Environmental Psychology*, *21*(3), 273–281.
- Hine, C. (2000). *Virtual ethnography*. Thousand Oaks, CA: Sage Publications Ltd.
- Howe, J. (2006). The Rise of Crowdsourcing. *Wired Magazine*, *14*(6), 1–4.

- Hutchins, E. (1995). *Cognition in the Wild*. Cambridge, MA: MIT Press.
- Ilggen, D., & Hollenbeck, J. (1991). The structure of work: Job design and roles. *Handbook of Industrial and Organizational Psychology*, 2, 165–207.
- Ilggen, D., Hollenbeck, J., Johnson, M., & Jundt, D. (2005). Teams in organizations: From Input-Process-Output Models to IMO models. *Annual Review of Psychology*, 56, 517–543.
- Intille, S., Larson, K., Beaudin, J., Nawyn, J., Tapia, E., & Kaushik, P. (2005). A living laboratory for the design and evaluation of ubiquitous computing technologies. In *Extended Abstracts on Human Factors in Computing Systems* (p. 1944). New York, NY.
- Irwin, A. (1995). *Citizen science: A study of people, expertise and sustainable development*. New York, NY: Routledge.
- Irwin, A. (2001). Constructing the scientific citizen: Science and democracy in the biosciences. *Public Understanding of Science*, 10(1), 1.
- Jensen, C., & Scacchi, W. (2005). Process modeling across the web information infrastructure. *Software Process: Improvement and Practice*, 10(3), 255–272.
- Kanefsky, B., Barlow, N., & Gulick, V. (2001). Can distributed volunteers accomplish massive data analysis tasks? *Lunar and Planetary Science*, 1.
- Kaufman, K. (1990). *A Field Guide to Advanced Birding: Birding Challenges and How to Approach Them*. New York, NY: Houghton Mifflin Harcourt.
- Kim, S., Robson, C., Zimmerman, T., Pierce, J., & Haber, E. (2011). Creek watch: Pairing usefulness and usability for successful citizen science. In *Proceedings of the 2011 International Conference on Human Factors in Computing Systems* (pp. 2125–2134). New York, NY.
- Kittur, A., & Kraut, R. (2008). Harnessing the wisdom of crowds in Wikipedia: Quality through coordination. In *Proceedings of the 2008 ACM conference on Computer supported cooperative work* (pp. 37–46). New York, NY.
- Klein, H., & Myers, M. (1999). A set of principles for conducting and evaluating interpretive field studies in information systems. *MIS Quarterly*, 67–93.
- Krasny, M., & Bonney, R. (2005). Environmental Education Through Citizen Science and Participatory Action Research. In *Environmental education and advocacy: Changing perspectives of ecology and education*. New York, NY: Cambridge University Press.
- Kruglinski, S., & Long, M. (2008). The 10 most influential people in science. *Discover Magazine*.

- Kuhn, T. (1970). *The Structure of Scientific Revolutions*. Chicago, IL: University of Chicago Press.
- Kumar, K. (1984). “People’s Science” and development theory. *Economic and Political Weekly*, 19(28), 1082–1084.
- Kvale, S., & Brinkmann, S. (2008). *InterViews: Learning the Craft of Qualitative Research Interviewing*. Thousand Oaks, CA: Sage Publications, Inc.
- Lave, J., & Wenger, E. (1991). *Situated Learning: Legitimate Peripheral Participation*. Cambridge University Press.
- Lawrence, A. (2006). ‘No Personal Motive?’ Volunteers, biodiversity, and the false dichotomies of participation. *Ethics, Place & Environment*, 9(3), 279–298.
- Lawrence, A. (2010). *Taking Stock of Nature: Participatory Biodiversity Assessment for Policy and Planning*. Cambridge, UK: Cambridge University Press.
- Lawrence, K., Finholt, T., & Kim, I. (2007). Warm fronts and high pressure systems: Overcoming geographic dispersion in a meteorological cyberinfrastructure project. In *Proceedings of the Hawai’i International Conference on Systems Sciences*. Washington, DC.
- Lee, C., Dourish, P., & Mark, G. (2006). The human infrastructure of cyberinfrastructure. In *Proceedings of the 2006 Conference on Computer supported cooperative work* (p. 492). New York, NY.
- Lincoln, Y., & Guba, E. (1985). *Naturalistic Inquiry*. Newbury Park, CA: Sage Publications, Inc.
- Logic Model Development Guide* (Tech. Rep.). (2004). Battle Creek, MI: W.K. Kellogg Foundation.
- Luhmann, N. (1995). *Social Systems*. Stanford University Press.
- Luke, D. (2005). Getting the big picture in community science: Methods that capture context. *American Journal of Community Psychology*, 35(3), 185–200.
- Luther, K., Counts, S., Stecher, K., Hoff, A., & Johns, P. (2009). Pathfinder: An online collaboration environment for citizen scientists. In *Proceedings of the 2009 international conference on human factors in computing systems* (pp. 239–248). New York, NY: ACM.
- March, J., Simon, H., & Guetzkow, H. (1958). *Organizations*. Wiley New York.
- Marks, M. A., Mathieu, J. E., & Zaccaro, S. J. (2001). A temporally based framework and taxonomy of team processes. *Academy of Management Review*, 26(3), 356–376.

- Markus, M., Manville, B., & Agres, C. (2000). What makes a virtual organization work? *Sloan Management Review*, 42(1), 13–26.
- Maron, N., Smith, K., & Loy, M. (2009). *Sustaining Digital Resources: An On-the-Ground View of Projects Today* (Tech. Rep.). New York, NY: Ithaka S+R.
- Maturana, H., & Varela, F. (1980). *Autopoiesis and Cognition: The Realization of the Living*. Springer.
- McAdam, D. (1999). *Political Process and the Development of Black Insurgency, 1930–1970*. University of Chicago Press.
- McAdam, D., & Scott, W. (2005). Organizations and movements. In *Social movements and organization theory* (pp. 4–40). New York, NY: Cambridge University Press.
- McCaffrey, R. (2005). Using citizen science in urban bird studies. *Urban Habitats*, 3(1), 70–86.
- McGinniss, M. (2011). An introduction to IAD and the language of the Ostrom workshop: A simple guide to a complex framework. *Policy Studies Journal*, 39(1), 169–183.
- McGrath, J. E., & Hollingshead, A. B. (1994). *Groups Interacting With Technology*. Thousand Oaks, CA: Sage Publications.
- Miller, S. (2001). Public understanding of science at the crossroads. *Public Understanding of Science*, 10(1), 115–120.
- Miller-Rushing, A., & Primack, R. (2008). Global warming and flowering times in Thoreau's Concord: A community perspective. *Ecology*, 89(2), 332–341.
- Millette, V., & Gagné, M. (2008). Designing volunteers' tasks to maximize motivation, satisfaction and performance: The impact of job characteristics on volunteer engagement. *Motivation and Emotion*, 32(1), 11–22.
- Mingers, J. (1995). *Self-Producing Systems: Implications and Applications of Autopoiesis*. Springer.
- Moore, E., & Koontz, T. (2003). A typology of collaborative watershed groups: Citizen-based, agency-based, and mixed partnerships. *Society and Natural Resources*, 16(5), 451–460.
- Niitamo, V., Kulkki, S., Eriksson, M., & Hribernik, K. (2006). State-of-the-art and good practice in the field of living labs. In *Proceedings of the 12th international conference on concurrent enterprising: Innovative products and services through collaborative networks* (pp. 26–28). Milan, Italy.

- Nov, O., Arazy, O., & Anderson, D. (2011). *Dusting for science: Motivation and participation of digital citizen science volunteers*. New York, NY: ACM.
- Oesterlund, C. (2008). Documents in place: Demarcating places for collaboration in healthcare settings. *Computer Supported Cooperative Work (CSCW)*, 17, 195-225.
- Olson, G., & Olson, J. (2000). Distance matters. *Human-Computer Interaction*, 15(2), 139-178.
- Olson, J. S., & Teasley, S. (1996). Groupware in the wild: Lessons learned from a year of virtual collocation. In *Proceedings of the 1996 ACM conference on Computer supported cooperative work* (pp. 419-427). New York, NY: ACM.
- Olson, M. (1965). *The Logic of Collective Action: Public Goods and the Theory of Groups*. Cambridge, MA: Harvard University Press.
- Orr, J. (1996). *Talking About Machines: An Ethnography of a Modern Job*. Ithaca, NY: ILR Press.
- Osborn, D., Pearse, J., & Roe, C. (2002). Monitoring rocky intertidal shorelines: A role for the public in resource management. In O. Magoon, H. Converse, B. Baird, B. Jines, & M. Miller-Henson (Eds.), *California and the World Ocean: Revisiting and Revising California's Ocean Agenda* (p. 624-636). Reston, VA: ASCE.
- Ostrom, E. (1986). An agenda for the study of institutions. *Public Choice*, 48(1), 3-25.
- Ostrom, E. (1990). *Governing the commons: The evolution of institutions for collective action*. New York, NY: Cambridge University Press.
- Overdeest, C., & Mayer, B. (2008). Harnessing the power of information through community monitoring: Insights from social science. *Texas Law Review*, 86, 1493-1526.
- Overdeest, C., Orr, C., & Stepenuck, K. (2004). Volunteer stream monitoring and local participation in natural resource issues. *Human Ecology Review*, 11(2), 177-185.
- Pearce, J. (1993). *Volunteers: The Organizational Behavior of Unpaid Workers*. New York, NY: Routledge.
- Perecman, E., & Curran, S. (2006). *A handbook for social science field research: Essays & bibliographic sources on research design and methods*. Thousand Oaks, CA: Sage Publications, Inc.
- Perlow, L. (1999). The time famine: Toward a sociology of work time. *Administrative Science Quarterly*, 44(1), 57-81.

- Phillips, T., Bonney, R., & Shirk, J. (2012). What is our impact? Toward a unified framework for evaluating educational impacts of citizen science. In J. Dickinson & R. Bonney (Eds.), *Citizen Science: Public Collaboration in Environmental Research*. Ithaca, NY: Cornell University Press.
- Powell, W., & DiMaggio, P. (1991). *The new institutionalism in organizational analysis*. Chicago, IL: University of Chicago Press.
- Powell, W., Koput, K., & Smith-Doerr, L. (1996). Interorganizational collaboration and the locus of innovation: Networks of learning in biotechnology. *Administrative Science Quarterly*, 116–145.
- Preece, J., & Shneiderman, B. (2009). The reader-to-leader framework: Motivating technology-mediated social participation. *AIS Transactions on Human-Computer Interaction*, 1(1), 13–32.
- Putnam, R. (2001). *Bowling Alone: The Collapse and Revival of American Community*. Simon and Schuster.
- Qin, J., Lancaster, F., & Allen, B. (1997). Types and levels of collaboration in interdisciplinary research in the sciences. *Journal of the American Society for Information Science and Technology*, 48(10), 893–916.
- Raddick, M., Bracey, G., Carney, K., Gyuk, G., Borne, K., Wallin, J., Jacoby, S., & Planetarium, A. (2009). Citizen science: Status and research directions for the coming decade. In *AGB Stars and Related Phenomena Astro 2010: The Astronomy and Astrophysics Decadal Survey* (Vol. 2010). Paris, France.
- Raddick, M., Bracey, G., Gay, P., Lintott, C., Murray, P., Schawinski, K., Szalay, A., & Vandenberg, J. (2010). GalaxyZoo: Exploring the motivations of citizen science volunteers. *Astronomy Education Review*, 9, 010103.
- Rao, H. (1998). Caveat emptor: The construction of nonprofit consumer watchdog organizations. *American Journal of Sociology*, 103(4), 912–961.
- Rettberg, S. (2005). All together now: Collective knowledge, collective narratives, and architectures of participation. In *Digital arts and culture*. Irvine, CA.
- Rojcewicz, R. (2006). *The Gods and Technology: A Reading of Heidegger*. Albany, NY: State University of New York Press.
- Roman, D. (2009). Crowdsourcing and the question of expertise. *Communications of the ACM*, 52(12), 12.

- Rotman, D., Preece, J., Hammock, J., Procita, K., Hansen, D., Parr, C., Lewis, D., & Jacobs, D. (2012). Dynamic changes in motivation in collaborative ecological citizen science projects. In *Proceedings of the ACM 2012 conference on Computer supported cooperative work*. New York, NY: ACM.
- Ruhleder, K. (2000). The virtual ethnographer: fieldwork in distributed electronic environments. *Field Methods*, 12(1), 3.
- Russell, E., Switzer, A., & Edelson, D. (2011). National Geographic FieldScope: A collaboratory geospatial platform for citizen science. In *Proceedings of Workshops at 2011 IEEE Seventh International Conference on eScience* (pp. 34–38). IEEE.
- Savan, B., Morgan, A., & Gore, C. (2003). Volunteer environmental monitoring and the role of the universities: The case of Citizens' Environment Watch. *Environmental Management*, 31(5), 561–568.
- Sawyer, S., & Eschenfelder, K. (2002). Social Informatics: Perspectives, examples, and trends. *Annual Review of Information Science and Technology*, 36, 427–65.
- Schmeller, D., Henry, P., Julliard, R., Gruber, B., Clobert, J., Dziock, F., Lengyel, S., Nowicki, P., Deri, E., Budrys, E., et al.. (2009). Advantages of volunteer-based biodiversity monitoring in Europe. *Conservation Biology*, 23(2), 307–316.
- Scott, W. (1991). Unpacking institutional arguments. In *The new institutionalism in organizational analysis* (pp. 143–63). Chicago, IL: University of Chicago Press.
- Scott, W., & Davis, G. (2007). *Organizations and Organizing: Rational, Natural, and Open System Perspectives*. Upper Saddle River, NJ: Pearson Prentice Hall.
- Shaw, T., & Jarvenpaa, S. (1997). Process models in information systems. In *Information Systems and Qualitative Research: Proceedings of the IFIP TC8 WG 8.2 International Conference on Information Systems and Qualitative Research* (p. 70). Boston, MA.
- Shepsle, K. (1989). Studying institutions. *Journal of Theoretical Politics*, 1(2), 131–147.
- Shirk, J., Ballard, H., Wilderman, C., Phillips, T., Wiggins, A., Jordan, R., McCallie, E., Minarcheck, M., Lewenstein, B., Krasny, M., & Bonney, R. (2012). Public participation in scientific research: A framework for deliberate design. *Ecology & Society*.
- Shirky, C. (2010). *Cognitive Surplus: Creativity and Generosity in a Connected Age*. ePenguin.
- Silvertown, J. (2009). A new dawn for citizen science. *Trends in Ecology & Evolution*, 24, 467–471.

- Sonnenwald, D. (2007). Scientific collaboration: A synthesis of challenges and strategies. *Annual Review of Information Science and Technology*, 41, 643–681.
- Spiro, M. (2004). What should the citizen know about science? *Journal of the Royal Society of Medicine*, 97(1), 50–50.
- Spradley, J. (1979). *The Ethnographic Interview*. Wadsworth Publishing Company.
- Spradley, J. (1980). *Participant observation*. Wadsworth Thomson Learning.
- Star, S., & Ruhleder, K. (1994). Steps towards an ecology of infrastructure: Complex problems in design and access for large-scale collaborative systems. In *Proceedings of the 1994 ACM Conference on Computer supported cooperative work* (pp. 253–264). New York, NY.
- Stinchcombe, A. (1965). Social structure and organizations. In J. March (Ed.), *The Handbook of Organizations* (pp. 142–193). Berkeley, CA: University of California Press.
- Stinchcombe, A. (1990). *Information and Organizations*. Berkeley, CA: University of California Press.
- Sturgis, P., & Allum, N. (2004). Science in society: Re-evaluating the deficit model of public attitudes. *Public Understanding of Science*, 13(1), 55–74.
- Sullivan, B., Wood, C., Iliff, M., Bonney, R., Fink, D., & Kelling, S. (2009). eBird: A citizen-based bird observation network in the biological sciences. *Biological Conservation*, 142(10), 2282–2292.
- Surowiecki, J. (2004). *The Wisdom of Crowds: Why the Many Are Smarter Than the Few and How Collective Wisdom Shapes Business, Economies, Societies, and Nations*. Doubleday Books.
- Tebes, J. (2005). Community science, philosophy of science, and the practice of research. *American Journal of Community Psychology*, 35(3), 213–230.
- The American Heritage Dictionary of the English Language, Fourth Edition*. (2010, Mar 26). <http://dictionary.reference.com/browse/participation>.
- The American Heritage Dictionary of the English Language, Fourth Edition*. (2012a, Mar 14). <http://dictionary.reference.com/browse/organizing>.
- The American Heritage Dictionary of the English Language, Fourth Edition*. (2012b, Apr 14). <http://dictionary.reference.com/browse/ground+truth>.
- Thompson, J., Zald, M., & Scott, W. (2003). *Organizations in Action: Social Science Bases of Administrative Theory*. Transaction Pub.

- Travis, J. (2008). Science and commerce: Science by the masses. *Science*, *319*(5871), 1750.
- Trochim, W., Marcus, S., Mâsse, L., Moser, R., & Weld, P. (2008). The evaluation of large research initiatives. *American Journal of Evaluation*, *29*(1), 8–28.
- Trumbull, D., Bonney, R., Bascom, D., & Cabral, A. (2000). Thinking scientifically during participation in a citizen-science project. *Science Education*, *84*(2), 265–275.
- Tushman, M., & Nadler, D. (1978). Information processing as an integrating concept in organizational design. *Academy of Management Review*, *3*(3), 613–624.
- Vaidyanathan, A., Krishnaji, N., & Kannan, K. P. (1979). People's Science movements. *Economic and Political Weekly*, *14*(2), 57–58.
- Viégas, F., Wattenberg, M., & McKeon, M. (2007). The hidden order of Wikipedia. *Lecture Notes in Computer Science*, *4564*, 445.
- Walsh, J., & Maloney, N. (2007). Collaboration structure, communication media, and problems in scientific work teams. *Journal of Computer-Mediated Communication*, *12*(2), 712–732.
- Walter, M. (1998). Participatory action research. *Reason*, 71.
- Wandersman, A. (2003). Community science: Bridging the gap between science and practice with community-centered models. *American Journal of Community Psychology*, *31*(3), 227–242.
- Wasko, M., Faraj, S., & Teigland, R. (2004). Collective action and knowledge contribution in electronic networks of practice. *Journal of the Association for Information Systems*, *5*(11-12), 493–513.
- Watson-Manheim, M., Chudoba, K., & Crowston, K. (2002). Discontinuities and continuities: A new way to understand virtual work. *Information Technology & People*, *15*(3), 191–209.
- Weick, K. (1969). *The Social Psychology of Organizing*. Reading, MA: Addison-Wesley.
- Weidensaul, S. (2007). *Of a Feather: A Brief History of American Birding*. Houghton Mifflin Harcourt.
- Weinstein, M. (2012). *Tams analyzer version 4.10b5h*. GPL licensed download from <http://tamsys.sourceforge.net>.
- Wenger, E. (1999). *Communities of Practice: Learning, Meaning, and Identity*. Cambridge University Press.

- Wiggins, A. (2010). Organizing from the middle out: Citizen science in the National Parks. In *Proceedings of the 2010 iConference*. Champaign, IL.
- Wiggins, A. (2011). eBirding: Technology adoption and the transformation of leisure into science. In *Proceedings of iconference 2011*. Seattle, WA.
- Wiggins, A., & Crowston, K. (2010). Developing a conceptual model of virtual organisations for citizen science [Research Note]. *International Journal of Organisational Design and Engineering*, 1, 148-162.
- Wiggins, A., & Crowston, K. (2012). Goals and tasks: Two typologies of citizen science projects. In *Proceedings of the Hawai'i International Conference on Systems Sciences (HICSS-45)*. Wailea, HI: IEEE.
- Wiggins, A., Newman, G., Stevenson, R., & Crowston, K. (2011). *Mechanisms for data quality and validation in citizen science*.
- Wilderman, C. (2004). From the field: A service providers experience with two operational models for community science. *Journal of Community Based Collaboratives Research*.
- Wilderman, C. C. (2007). Models of community science: Design lessons from the field. In C. McEver, R. Bonney, J. Dickinson, S. Kelling, K. Rosenberg, & J. L. Shirk (Eds.), *Citizen science toolkit conference*. Cornell Laboratory of Ornithology, Ithaca, NY.
- Wilkinson, D. (2008). Strong regularities in online peer production. In *Proceedings of the 9th ACM Conference on Electronic Commerce* (pp. 302–309). New York, NY.
- Williams, D., & Vaske, J. (2002). The measurement of place attachment: Validity and generalizability of a psychometric approach. *Forest Science*, 49, 830–840.
- Williams, R. (1995). Constructing the public good: Social movements and cultural resources. *Social Problems*, 42, 124.
- Williamson, O. (1981). The economics of organization: The transaction cost approach. *American Journal of Sociology*, 548–577.
- Wilson, J. (2000). Volunteering. *Annual Review of Sociology*, 26, 215–240.
- Wivell, T. (2011). *Passport to AMC's High Huts in the White Mountains*. Boston, MA: Appalachian Mountain Club Books.
- Wood, C., Sullivan, B., Iliff, M., Fink, D., & Kelling, S. (2011). eBird: Engaging birders in science and conservation. *PLoS Biology*, 9(12).
- Wulf, W. (1993). The collaboratory opportunity. *Science*, 261(5123), 854.

Yin, R. (1984). *Case Study Research: Design and Methods*. Thousand Oaks, CA: Sage Publications, Inc.

CURRICULUM VITAE

ANDREA WIGGINS

PhD Candidate

Syracuse University

School of Information Studies

337 Hinds Hall

Syracuse, NY 13244 USA

Cell: +1 315-558-0122

Email: awiggins@syr.edu

Web: <http://www.andreawiggins.com>

Research Interests

LARGE-SCALE COLLABORATION

My research investigates the interactions of organizational settings, work design, and technology use. My interests include data-intensive science, distributed collaboration, and social computing. The focus of my dissertation is on the role of technologies in supporting public participation in scientific research.

Education

SYRACUSE UNIVERSITY SCHOOL OF INFORMATION STUDIES Syracuse, NY

PhD, Information Science & Technology, 2012

DISSERTATION: Crowdsourcing Scientific Work: A Comparative Study of
Technologies, Processes, and Outcomes in Citizen Science

Committee: Kevin Crowston (Advisor), Steve Sawyer, Jian Qin, Rick Bonney, Geoffrey Bowker, and Murali Venkatesh.

UNIVERSITY OF MICHIGAN SCHOOL OF INFORMATION Ann Arbor, MI

MSI, Information, 2007 (*ALA accredited*)

THESIS: Exploring Peer Prestige in Academic Hiring Networks

Committee: Michael J. McQuaid (Advisor), and Lada A. Adamic.

ALMA COLLEGE Alma, MI

BA, Mathematics, 2000

Publications

REFEREED JOURNAL PUBLICATIONS (5)

Shirk, J., Ballard, H., Wilderman, C., Phillips, T., Wiggins, A., Jordan, R., McCallie, E., Minarchek, M., Lewenstein, B., Krasny, M., & Bonney, R. (In Press). Public Participation in Scientific Research: A Framework for Deliberate Design. *Ecology & Society*.

Crowston, K., Wei, L., Howison, J., & Wiggins, A. (2012). Free/Libre Open Source Software Development: What We Know and What We Do Not Know. *ACM Computing Surveys*, 44(2).

Wiggins, A. & Sawyer, S. (2012). Intellectual Diversity and the Faculty Composition of iSchools. *Journal of the American Society for Information Science and Technology*. 63(1), 8–21.

Howison, J., Wiggins, A., & Crowston, K. (2011). Validity Issues in the Use of Social Network Analysis for the Study of Online Communities. *Journal of the Association for Information Systems*, 12(12).

Wiggins, A. & Crowston, K. (2010). Developing a Conceptual Model of Virtual Organizations for Citizen Science. *International Journal of Organisational Design and Engineering*, 1(1/2), 148–162.

REFEREED CONFERENCE PAPERS (15)

Wiggins, A. & Crowston, K. *Goals and Tasks: Two Typologies of Citizen Science Projects*. Proceedings of the 45th Annual Hawaii International Conference on System Sciences. Wailea, HI, 4–7 January, 2012.

- Wiggins, A. & Crowston, K. *From Conservation to Crowdsourcing: A Typology of Citizen Science*. Proceedings of the 44th Annual Hawaii International Conference on System Sciences. Koloa, HI, 4–7 January, 2011.
- Qin, J., Chen, M., Liu, X. & Wiggins, A. *Linking Entities in Scientific Metadata*. Proceedings of the International Conference on Dublin Core and Metadata Applications. Pittsburgh, PA, 20–22 October, 2010.
- Wiggins, A. & Crowston, K. *Reclassifying Success and Tragedy in FLOSS Projects*. Proceedings of the Sixth International Conference on Open Source Software. Notre Dame, IN, 30 May–2 June, 2010.
- Wiggins, A. & Sawyer, S. *Intellectual Diversity in iSchools: Past, Present and Future*. Proceedings of the iConference 2010. Champaign, IL, 3–6 February, 2010.
- Crowston, K., Wiggins, A., & Howison, J. *Analyzing leadership dynamics in distributed group communication*. In Proceedings of the 43rd Annual Hawaii International Conference on System Sciences. Kauai, HI, 5–8 January, 2010.
- Wiggins, A., Howison, J. & Crowston, K. *Heartbeat: Measuring Active User Base and Potential User Interest in FLOSS Projects*. Proceedings of the Fifth International Conference on Open Source Software. Skövde, SE, 3–6 June, 2009.
- Crowston, K., Wiggins, A. & Howison, J. *Opportunities for eScience Research on Free/Libre Open Source Software*. Proceedings of the Oxford e-Research Conference 2008. Oxford, UK, 11–13 September, 2008.
- Wiggins, A., Howison, J., & Crowston, K. *Replication of FLOSS Research as eResearch*. Proceedings of the Oxford e-Research Conference 2008. Oxford, UK, 11–13 September, 2008.
- Wiggins, A., Howison, J. & Crowston, K. *Social dynamics of FLOSS team communication across channels*. Proceedings of the Fourth International Conference on Open Source Software. Milan, IT, 7–10 September, 2008.
- Howison, J., Wiggins, A. & Crowston, K. *eResearch workflows for studying free and open source software development*. Proceedings of the Fourth International Conference on Open Source Software. Milan, IT, 7–10 September, 2008.
- Crowston, K., Howison, J. & Wiggins, A. *eSocial Science for Free/Libre Open Source Software Researchers*. Proceedings of the Fourth International Conference on eSocial Science. Manchester, UK, 18–20 June, 2008.
- Wiggins, A., McQuaid, M., & Adamic, L. *Community Identity: Peer Prestige and Academic Hiring in the iSchools*. iConference 2008. Los Angeles, CA, 28 February–1 March, 2008.

Wiggins, A. *Data Driven Design: Using Web Analytics to Improve Information Architectures*. Proceedings of the ASIST Information Architecture Summit 2007. Las Vegas, NV, 22–26 March, 2007.

Wiggins, A., Adamic, L., & McQuaid, M. *The Small Worlds of Academic Hiring Networks*. Applications of Social Network Analysis 2006. Zürich, CH, 5–6 October, 2006.

REFEREED CONFERENCE POSTERS (10)

Wiggins, A. *eBirding: Technology Adoption and the Transformation of Leisure into Science*. iConference 2011. Seattle, WA, 8–11 February, 2011. *Best Poster Award*.

Qin, J., D'Ignazio, J., Oakleaf, M., Crowston, K., Wiggins, A. & Steinhart, G. *Educating eScience Librarians*. Sixth International Digital Curation Conference. Chicago, IL, 6–8 December, 2010.

Wiggins, A. *Crowdsourcing Science: Organizing Virtual Participation in Knowledge Production*. GROUP 2010. Sanibel Island, FL, 7–11 November, 2010.

Wiggins, A. *Technology and Work Practices in Citizen Science*. ASIST 2010. Pittsburgh, PA, 22–27 October, 2010.

Wiggins, A. *Organizing from the Middle Out: Citizen Science in the National Parks*. iConference 2010. Champaign, IL, 3–6 February, 2010.

Wiggins, A. *Coordination Dynamics in Free/Libre Open Source Software Development*. In Proceedings of the Doctoral Consortium of the Fifth International Conference on Open Source Software. Skövde, SE, 3–6 June, 2009.

Crowston, K., Bolici, F., & Wiggins, A. *Appropriate Measures for On-line Communities*. Poster presented at Sunbelt XXIX. San Diego, CA, 10–15 March, 2009.

Wiggins, A. *Interdisciplinary Diversity in the iSchool Community*. iConference 2009. Chapel Hill, NC, 9–11 February, 2009.

Li, Q., Heckman, R., Allen, E., Crowston, K., Eseryel, U., Howison, J., & Wiggins, A. *Asynchronous Decision-Making in Distributed Teams*. CSCW 2008. San Diego, CA, 8–12 November, 2008.

Williams, K., Bishop, A., Bard, M., & Wiggins, A. *Turning program management data into a research dataset: working with the Department of Commerce to study technology projects in local communities*. IASSIST 2006. Ann Arbor, MI, 23–26 May, 2006.

PANEL PRESENTATIONS (3)

Yew, J., Chi, E., Antin, J., Paul, S., Shaw, A., Lampe, C., & Wiggins, A. *Apples to Oranges?: Comparing across studies of open collaboration/peer production*. Panel at the 2011 International Symposium on Wikis and Open Collaboration. Mountain View, CA, 5 October, 2011.

Khoo, M., Twidale, M., Wiggins, A., & Rode, J. *The Sociotechnical Learning Object Slam Happening*. Alternative event at the 2011 iConference. Seattle, WA, 8–11 February, 2011.

Crowston, K., Sawyer, S., Ribes, D., & Wiggins, A. *Big eScience, Little eScience*. Panel presentation at iConference 2009. Chapel Hill, NC, 9–11 February, 2009.

OTHER CONFERENCE AND WORKSHOP PRESENTATIONS (7)

Wiggins, A., Newman, G., Stevenson, R. D., & Crowston, K. *Mechanisms for Data Quality and Validation in Citizen Science*. Paper presented at “Computing for Citizen Science” workshop at the IEEE eScience Conference. Stockholm, SE, 5 December, 2011.

Shirk, J., Ballard, H., Wiggins, A., Jordan, R., & Bonney, R. *Public participation in scientific research: A framework for intentional design*. Presentation at the 96th Ecological Society of America Annual Meeting. Austin, TX, 7–12 August, 2011.

Wiggins, A. & Crowston, K. *Distributed Scientific Collaboration: Research Opportunities in Citizen Science*. Paper presented at “The Changing Dynamics of Scientific Collaboration” Workshop at CSCW 2010. Savannah, GA, 7 February, 2010.

Wiggins, A. & Crowston, K. *Designing Virtual Organizations for Citizen Science*. Paper presented at OASIS 2009 Workshop. Phoenix, AZ, 15 December, 2009.

Wiggins, A. *Tales of the Field: Building Small Science Cyberinfrastructure*. Presentation at Society for Social Studies of Science 2009 Annual Meeting 4S 2009. Washington, DC, 28 October–1 November, 2009.

Wiggins, A., Crowston, K., & Howison, J. *Collaborative Data Analysis with Taverna Workflows*. Paper presented at Microsoft eScience 2009 Workshop. Pittsburgh, PA, 15–17 October, 2009.

Crowston, K., Howison, J., & Wiggins, A. *eScience to support research on free/libre/open source software development*. Poster presented at Microsoft eScience 2009 Workshop. Pittsburgh, PA, 15–17 October, 2009.

INVITED PRESENTATIONS (10)

- Wiggins, A. *Motivation by Design: Technologies, Experiences, and Incentives*. Presentation at The Citizen Cyberscience Summit 2012. London, UK, 16 February, 2012.
- Wiggins, A. *Secondary data analysis with digital trace data: examples from FLOSS research*. Methodology seminar presentation for Département LUSSI, Institut Télécom Bretagne. Technopole Brest-Iroise, FR, 13 July, 2011.
- Wiggins, A. *Open Source, Open Science, and Citizen Science*. Research seminar presentation for Département LUSSI, Institut Télécom Bretagne. Technopole Brest-Iroise, FR, 24 June, 2011.
- Goggins, S., & Wiggins, A. *Research methodologies for digital trace data*. Roundtable discussion at the University of British Columbia, School of Library, Archival and Information Sciences. Vancouver, BC, 4 February, 2011.
- Wiggins, A. *Little eScience*. Lecture for Information Management Group, University of Manchester. Manchester, UK, 18 June, 2009.
- Wiggins, A. *Dynamic Social Network Analysis (and more!) with eResearch Workflows*. Presentation at OSSWatch Expert Workshop on Profiling Communities. Oxford, UK, 21 July, 2008.
- Wiggins, A. *Data Analysis*. Guest lecture for IST 400/600, Science Data Management, Syracuse University. Syracuse, NY, 14 April, 2008.
- Conklin, M., Howison, J., & Wiggins, A. *FLOSSmole*. Presentation at NSF Free and Open Source Software Research Infrastructure Workshop. Irvine, CA, 11–12 February, 2008.
- Wiggins, A. *Web Analytics*. Guest lecture for SI 658, Information Architecture, University of Michigan. Ann Arbor, MI, 10 October, 2006.
- Wiggins, A., Bard, M., & Baugh, J. *Dabbling in Drupal: A Case Study of the University of Michigan Community Information Corps Website*. Presentation for the Library Technologies Division at the Michigan Library Association annual conference. Grand Rapids, MI, 25 October, 2005.

OTHER PUBLICATIONS (4)

- Wiggins, A. *Building a Data-Backed Persona*. Boxes & Arrows, 2007.
<http://www.boxesandarrows.com/view/building-a-data>
- Wiggins, A. *Data Driven Design: Using Web Analytics to Validate Heuristics*. Bulletin of the American Society for Information Science and Technology, 2007.
<http://www.asis.org/Bulletin/Jun-07/wiggins.html>

Wiggins, A. *Metrics for Heuristics: Quantifying User Experience Part 2*. Boxes & Arrows, 2006. http://www.boxesandarrows.com/view/metrics_for_heu7

Wiggins, A. *Metrics for Heuristics: Quantifying User Experience Part 1*. Boxes & Arrows, 2006. http://www.boxesandarrows.com/view/metrics_for_heu

WORK UNDER REVIEW (1)

Newman, G., Wiggins, A., Crall, A., Graham, E., Newman, S., & Crowston, K. Citizen Science Futures: Emerging Technologies and Shifting Paradigms. *Frontiers in Ecology and the Environment*.

Research Experience

SYRACUSE UNIVERSITY, SCHOOL OF INFORMATION STUDIES

Research Assistant to Kevin Crowston, 2011–2012.

NSF IIS Grant 09-68470. Community liaison for citizen science game development project.

Research Assistant to Kevin Crowston, 2009–2011.

NSF OCI grant 09-43049. Developed typology of citizen science projects; conducted online survey; case study data collection and analysis; workshop coordination.

Graduate Assistant to Kevin Crowston, 2008–2009.

Developed research infrastructure; grantwriting; DOI assignment.

Research Assistant to Kevin Crowston, 2008.

NSF CRI grant 07-08437. Developed analysis workflows and open access repository.

Graduate Assistant to Renee Franklin, 2007–2008.

Responsible for course management activities and research support.

Grantwriting, 2008–2010.

SoCS: Socially Intelligent Computing to Support Citizen Science: NSF IIS Grant 09-68470, \$478,858 for 36 months, awarded September 2010. PI: Kevin Crowston.

VOSS: Theory and Design of Virtual Organizations for Citizen Science: NSF OCI Grant 09-43049, \$150,000 for 24 months, awarded September 2009. PI: Kevin Crowston.

VOSS: Cultivating Effective Virtual Organizations: Unfunded NSF grant proposal, 2008.

INSTITUT TÉLÉCOM BRETAGNE

Visiting Scholar with Département de Logique des usages, sciences sociales et de l'information (LUSI), June–July 2011. Research exchange and collaboration, focusing on participation in online communities of practice, working with Nicolas Jullien.

UNIVERSITY OF MANCHESTER, SCHOOL OF COMPUTER SCIENCES

Project/Area Liaison for myGrid, 2009–2011. Ongoing collaboration and invited participation in UK eScience Review (December 2009). Project representative for NSF focus group on sustaining science gateways (June 2010).

Academic Visitor with myGrid at University of Manchester School of Computer Sciences, June 2009. Collaboration on scientific computing infrastructure development, supported by EPSRC grant EP/G026238/1, “myGrid: A Platform for e-Biology Renewal.”

UNIVERSITY OF MICHIGAN, SCHOOL OF INFORMATION

Data Archives Intern for the Technology Opportunities Project Data Archive project, January–August 2006. Evaluated, analyzed, and documented data archives for repository deposit.

Teaching Experience

SYRACUSE UNIVERSITY SCHOOL OF INFORMATION STUDIES

IST 600, Workflows: eScience and eResearch, 2011

Instructor: Developed and delivered a new one-credit graduate course providing a hands-on introduction to data-intensive research tools.

IST 777, Statistical Methods in Information Science and Technology, 2008

Practicum: Assisted in course redevelopment for modular hybrid delivery, including text selection and evaluation design.

IST 500, Distributed Collaboration and Emerging Technologies, 2008

Practicum: Assisted in new course design, text selection, and syllabus development.

IST 400/600, Science Data Management, 2008

Practicum: Developed and evaluated quizzes; graded papers; developed and delivered a lesson on data analysis.

IST 668, Literacy Through School Libraries, 2008

Assistant: Monitored online class discussions; handled student questions; graded papers.

IST 972, School Media Practicum, 2007–2008

Assistant: Maintained records for NY School Library Media certification requirements.

IST 335, Introduction to the Information-Based Organization, 2007

Practicum: Developed new course materials; adapted and lead in-class simulation exercise.

UNIVERSITY OF BRITISH COLUMBIA

Web Analytics for Site Optimization, 2006

Associate Instructor: Content contributor and editor for online course in Award of Achievement in Web Analytics series.

Professional Activities

HONORS & AWARDS

Best Poster Award, iConference 2011. One of three best posters out of 100 refereed presentations.

Phi Kappa Phi, Syracuse University, 2008.

School of Information Studies Fellowship, Syracuse University, 2007.

Stephen Markel Award, University of Michigan School of Information, 2006. Selected by Dean for “an extraordinary combination of talent, skill, dedication, leadership, compassion, humor, and entrepreneurial spirit in educational, professional, and personal pursuits.”

Segal AmeriCorps Education Award, National Service Trust, 2001. Award recognizing national service through AmeriCorps.

DOCTORAL CONSORTIA, COLLOQUIA, AND SYMPOSIA

Social Computational Systems. Minneapolis, MN, 9–11 June, 2011.

International Conference on Information Systems. St. Louis, MO, 8–11 December, 2010.

GROUP 2010, ACM International Conference on Supporting Group Work. Sanibel Island, FL, 7–11 November, 2010.

Organizational Communication and Information Systems Division, Academy of Management. Montréal, QC, 6–7 August, 2010.

Consortium for the Science of Sociotechnical Systems, Summer Research Institute.
Stevenson, WA, 13–17 June, 2010.

Fifth International Conference on Open Source Software. Skövde, SE, 3–6 June, 2009.

iConference 2009. Chapel Hill, NC, 9–11 February, 2009.

MEMBERSHIPS

Association for Computing Machinery and SIGCHI, since 2005

Ecological Society of America, since 2010

Service

COMMITTEE SERVICE

Syracuse University School of Information Studies

Personnel Committee (5 cases), 2010–2011

Doctoral Program Committee, 2008–2010

PhD Student Mentor, 2008–2009

Faculty Search Committee (8 positions), 2008

PhD Faculty Representative, 2007–2008

Advisory

Smithsonian Institution National Museum of Natural History, BioCube, April 2012

Adler Planetarium, NSF DRL grant 09-17608, “Investigating Audience Engagement with Citizen Science,” June 2011

Working Groups

Co-Chair, DataONE Working Group on Public Participation in Scientific Research,
2011–2016

Member, Northeast Regional Phenology Network Citizen Science working group,
2009–2011

CONFERENCE POSITIONS

Program Committees

iConference 2011

iConference 2006

WikiSym 2012

Workshop Organizer

Data Intensive Collaboration in Science and Engineering, with M. Bietz, C. Aragon, & M. Handel. CSCW 2012, Bellevue, WA, 11 February, 2012.

Design for Citizen Science, with K. Crowston. Blue Mountain Lake, NY, 24–27 May, 2011.

REVIEWING

Journals

CSCW Journal, 2012

Trends in Ecology & Evolution, 2011

Conferences

2012: CSCW, WikiSym

2011: CSCW, iConference, CHI, HICSS

2010: ASIST, CSCW, EurAM, HICSS, ICIS, OSS

2009: OSS, HICSS

2008: ICWSM

Other Activities

EMPLOYMENT EXPERIENCE

2005–2007: *Enlighten* Ann Arbor, MI — Data Analyst

2002–2005: *The Purple Rose Theatre Company* Chelsea, MI — Office Manager

2001–2002: *Nonprofit Enterprise at Work* Ann Arbor, MI — Education Program Associate

2000–2001: *Washtenaw Literacy* Ypsilanti, MI — Volunteer Coordinator

JURIED EXHIBITIONS

2010 *Emilee 09* (March 25, 2010). Silver gelatin print. Matrilineage: Sensational Remix. Spark Contemporary Art Space. Syracuse, NY.

2007 *Links* (March - April 2007). Silver gelatin print. Cheap Shots: Silver Dreams of Plastic Cameras. Gallery 4. Ann Arbor, MI.

2007 *iCommunity* (January - February 2007). Large format lightjet print. If Other Please Explain. WORK Gallery. Ann Arbor, MI.