EXPLORING COLLECTIVE INTELLIGENCE GAMES WITH DESIGN SCIENCE: A CITIZEN SCIENCE DESIGN CASE

Nathan Prestopnik

Syracuse University
343 Hinds Hall
Syracuse, NY 13244, USA
napresto@syr.edu

Kevin Crowston

Syracuse University
343 Hinds Hall
Syracuse, NY 13244, USA
crowston@syr.edu

ABSTRACT

Citizen science is a form of collective intelligence where members of the public are recruited to contribute to scientific investigations. Citizen science projects often use web-based systems to support collaborative scientific activities, but finding ways to attract participants and confirm the veracity of the data produced by non-scientists are key research questions. We describe a series of web-based tools and games currently under development to support taxonomic classification of organisms in photographs collected by citizen science projects. In the design science tradition, the systems are purpose-built to test hypotheses about participant motivation and techniques for ensuring data quality. Findings from preliminary evaluation and the design process itself are discussed.

INTRODUCTION

Citizen science is a phenomenon where members of the public are recruited to contribute to scientific investigations [4, 29]. Citizen science projects include asking participants to help classify astronomical photographs, report bird sightings, count insects in the field, or use spatial reasoning skills to align genomes or fold protein strings. Such activities draw many individuals into a cooperative endeavor toward a common scientific goal. They feature, "groups of individuals doing things collectively that seem intelligent," [17]. As such, citizen science may be considered a form of collective intelligence.

The citizen science phenomenon is relatively new, and a variety of open questions are of great interest to information science researchers and scientists from other fields who may wish to use citizen science approaches in their own research. An interesting and sometimes challenging issue for citizen science is that some scientific topics are highly "charismatic" but many others are not. For example, bird watching, astronomy, and conservation all have a certain cachet, even for non-enthusiasts. However, important work is also being conducted in areas that attract much less public interest, such as moth, mold, or lichen classification. While enthusiasts exist for virtually all areas of the natural sciences, collective intelligence systems rely on large numbers of participants, and how charismatic a branch of science is sometimes determines how well it can be systematized into a collective effort. As a result, the motivations of citizen science participants are important to understand, to attract new participants and retain old ones.

The quality of data produced by non-expert citizens when participating in collective scientific activities is a second topic of concern, as are techniques to turn science into tasks that can be performed by novices. The specific interest of this research, therefore, is to explore the relationships that exist between citizen science/collective intelligence system design, attraction and retention of participants, and the impact of these on data quality.

Unfortunately, it is difficult to use current, real-world citizen science projects as vehicles for exploring motivation, participation, users, technology, and data quality. Most challenges are practical: citizen science project developers, researchers, and managers have little time available to devote toward research projects not directly related to their specific object of inquiry. Because currently instantiated citizen science projects are working production systems, it is difficult to adjust project parameters, conduct experiments, issue surveys, interview participants, or otherwise gather information about the citizen science phenomenon. Invasive data collection efforts are likely to be disruptive and may have deleterious impacts on existing participant enthusiasm and data quality. In short, the potential drawbacks of granting complete access to the information science researcher vastly outweigh any benefits that might accrue.

On the other hand, low-impact methods of investigation (e.g., interviewing or surveying staff members or researchers, passively gathering information about project websites and systems, etc.) are less likely to produce data required to address motivational and data-quality questions. Studying citizen science without fine control over the systems of interest creates a different problem: artificiality will infect any knowledge generated by such research, as simulations, mock-ups, and de-contextualized inquiry substitute for realistic exploration of actual systems that are highly situated within complex problem spaces.

We address these challenges by developing systems explicitly designed to serve a dual purpose as vehicles for scientific inquiry and as functional and useful systems built and deployed to solve specific, real-world problems. Building systems is not a new approach to research, but the approach has recently been reconceptualized under the name design science. This approach resides in the familiar territory of system design and evaluation, but wraps these well-known activities around a broader research agenda targeted at natural or social-psychological science. The strength of this approach is that complex phenomenon such as collective intelligence and/or citizen science can be explored in a very realistic manner, while maintaining a great deal of control.

This paper is divided into two parts. First, a discussion of design science is presented. Second, an ongoing design science project in the collective intelligence and citizen science domains is described. This project involves the creation of several games and tools to support an important science task in the biological sciences: species classification. Results from the design process so far and from preliminary evaluations are reported. Discussions of the design science approach as a vehicle for collective intelligence scholarship follow.

PART I: DESIGN SCIENCE

Design Science Overview

Design science is an approach to scholarly study that couples traditional research methodologies with the development of an IT artifact to address natural science or social-psychological research questions coupled with design-related problems [11, 18, 19]. Design science is practiced (mostly without using the term) in many domains, particularly human-computer interaction (HCI) and computer science (CS) more generally, where innovative system design is common. The term and its formal conceptualization come from the field of information systems (IS), where system design is often viewed as atheoretical and so not research. In this setting, rigorous conceptualizations of design as a research tool are necessary to encourage its broader acceptance. However, even in fields where system design is embraced, the reconceptualization can be valuable, as the focus on designing useful artifacts often results in inattention to larger research questions. For example, in [7], many HCI evaluation practices are criticized as "usability evaluations" instead of scientific "evaluations for research", what [8] calls the "I did this and it's cool" form of study.

Design science research has two equally important outcomes: 1) a functional IT artifact that helps address a specific, challenging, and practical design problem within a given context, and 2) meaningful scholarly contributions to a field of inquiry. Compared to typical social-science research approaches, the design science approach requires additional components, including interactions with subject-matter experts (SMEs), a situational focus on the context in which a design will be deployed as well as system building and testing. Compared to typical systems research, the approach requires explicit use of theory to guide design decisions and—importantly—an ability to draw more general conclusions about these theories. In many ways, design science is like case study research, but featuring design cases that have been rigorously defined, designed, and developed as part of the research process [24].

The problem spaces addressed by design science inquiry are typically complex, sometimes referred to as "wicked" problems because they defy easy or obvious answers [2, 3, 23]. Problems suitable for a design science approach include both those that are unsolved and those which offer opportunities for newer or better solutions [11]. However, to be meaningful to researchers outside of the specific problem space, the IT artifact must also become a vehicle for broader natural science or social-psychological inquiry. Theory, design and evaluation are thus interrelated in design science research, coherent pieces of a whole [20] and conducted iteratively [11, 18].

Theory: The word "theory" is used broadly here [10], encompassing the adoption of existing theory as a lens through which to approach design, as well as consultation with experts and review of non-theoretical, project-specific design literature. This stage may also result in the generation of new theory, produced either from literature or from data, and conceptualized either prior to design of the IT artifact, during its development, or after its evaluation. The theory stage may be seen as both a beginning and an end to design science research: theory adopted early will inform design, and new theory will come from it.

Design: Design science research revolves around the design of an IT artifact, where theoretical and practical underpinnings shape a functional system. The designed artifact may ultimately produce new theory, so artifact design must take future evaluation into account. The design scientist must always keep in mind the research questions to be addressed through research evaluation of the artifact.

Evaluation: The evaluation stage is about more than saying "yes this worked," or, "no, this didn't work." It must address the project's broader research questions by validating adopted theory or leading to the generation of new theory. Evaluation is not always an end point for research; evaluation will often suggest ways to improve the artifact (as a system to address the problem space or as a research tool) in its next design iteration.

PART II: CITIZEN SCIENCE DESIGN CASE

Having described the nature of design science in the abstract, a concrete example will be of value. In this section, a collective intelligence project situated in the citizen science domain is described in detail, with emphasis on research goals, the problem space, and its design parameters.

Research Goals

Our study addresses two research questions. First, a critical issue in collective intelligence systems in general, and citizen science systems in particular, is attracting and retaining enough participants to make achievement of project goals possible. Systems with too little participation will be unlikely to generate meaningful quantities of scientific data.

To address this question, we draw on psychological theories about motivation [e.g. 5]. In [17], three basic motivations for individuals who are engaged in collective intelligence activities are suggested: money, love, and glory. For citizen science projects, offering payment to participants is rarely an option (project resources are typically too low), and most participants do not expect compensation for their efforts. Instead, participants indicate that inherent interest in the subject of scientific inquiry, the relevance of data collection efforts to particular interests or hobbies, the perception that a project will be fun and engaging, an interest in collaborate with experts, altruistic reasons, and hope for broader recognition as reasons for becoming involved in citizen science projects [1, 12, 21, 22, 28]. These reasons match well with the notions of "love" and "glory" as motivators [17]. There has been less scholarly or practical attention paid to how citizen science systems might be designed to motivate participants who do not hold these predominantly intrinsic motivations. As a result, most citizen science projects rely heavily on participants who have preexisting enthusiasm for the scientific topic of the project, be it astronomy, bird watching, or classifying insects.

In the broader collective intelligence domain, several models for attracting participation have been deployed. In systems such as von Ahn's reCAPTCHA [27], the collective intelligence system is established as an obstacle between users and their goals; reCAPTCHAs are used to verify that login attempts to web systems are coming from a human user, and to log in, users must use the reCAPTCHA tool. Other systems, such as the ESP game (an image tagging system) [26], Phetch (which produces accessible descriptions of images) [25], or TagATune (where users tag music clips) [13] are designed as games, capitalizing on "love" forms of motivation, and giving people enjoyable activities to undertake while also producing meaningful work almost as a by-product.

Games in particular seem to have great potential as a motivator for participation and as a tool for producing high quality scientific data. However, from a review of citizen science websites [29], it seems that few existing projects use games to motivate participation. Notable exceptions include *Fold It*, which disguises the science of protein string folding as a highly engaging puzzle game, and *Phylo*, where players compare genetic sequences in a colorful and abstract puzzle game. Both capitalize on human spatial reasoning abilities. The *Fold It* player pages (http://fold.it/portal/players) reveals that more than 300,000 players are contributing to this project; furthermore, *Fold It* recently made headlines for an important AIDS research breakthrough generated by players of the game. Some projects, like *Stardust@Home*, incorporate game-like elements such as leader boards, high scores, or other participation metrics, but do not frame their scientific activities as games per se. Scholarly study of collective intelligence games and games for citizen science may produce insights into how different participant groups can be attracted to citizen science projects and motivated to participate in them.

Our second research question is about techniques for ensuring data quality, a necessary precondition for further scientific use of the data, but difficult for several reasons. First, for many scientific problems there is "ground truth" of correct answers. Participants opinions are not inherently valid as they might be in systems designed to produce, for example, image tags for search engines. For data to be scientific, valid, and accepted, the right answers must be produced by participants and confirmed by experts. Second, in many areas of science, specialized knowledge is required to provide data, but few citizen science participants are experts. Furthermore, the effect of systems (especially game-like interactions) on data quality is largely unknown. Therefore, finding methods to turn scientific tasks into things that non-scientists can do well, as well as finding techniques to confirm the validity of participant-provided data, are important research goals. To address these questions, we draw on theories from the problem domain, which we describe next.

Problem Space

The problem space we address in this research comes from the biological sciences, particularly entomology, botany, and oceanography. In this domain, experts, enthusiasts, and curious members of the general public routinely collect and upload photographs of different living things. A photograph of an insect, plant, or animal, tagged with the date and location where it was taken, can provide valuable scientific data, e.g., on how urban sprawl impacts local ecosystems or evidence of local, regional, or global climactic shifts. However, to be useful, it is necessary to know what the picture is of, expressed in scientific terms, i.e., the scientific name of the species depicted. Some participants have the necessary knowledge (e.g., avid birders can identify particular bird species), but many potential participants do not.

To identify the species of specimens, biologists have developed taxonomic keys, which identify species from their particular combinations of characteristics, known as character-state combinations (i.e., attributes and values). The specific characters and states vary by taxon, but are broadly similar in structure. For example, a moth character might be its “orbicular spot,” with states including, “absent,” “dark,” “light,” etc. Given sufficient characters and states, it is possible to identify a photographed specimen to a specific family, genus, species, or even sub-species.

A challenging aspect of this problem is that researchers working within the same biological or ecological disciplines do not necessarily agree upon taxonomic keys. In fact, many researchers develop their own key variations to support their own specific research endeavors. Keys are therefore typically written for expert users, and are often complex, highly variable, and difficult to translate into a form that will be suitable for use in a collective intelligence system, where expert understanding of characters, states, and taxonomic identification cannot be assumed.

A second challenge is that even with an established key, some characters and states are beyond the ability of members of the general public to identify without training (e.g. the previous “orbicular spot” example). Others require true expert knowledge to apply (for example, classifying species by their sex organs). In some cases, especially for sub-species, true identifications cannot be made without access to specialized equipment; for example, some species are distinguishable only through their genetic makeup. This means that an IT artifact designed to support the classification task will be unlikely to effectively support both extremely knowledgeable users and extremely novice users; experts will require advanced tools with great flexibility, while novices may require simplified systems that have expert knowledge pre-built into them. In both cases, a web-based classification system will only be able to support some kinds of characters and states, while others will be impossible.

Design Parameters

To explore the motivations of citizen science participants and address the challenge of species classification in the biological sciences, a series of IT artifacts were designed and implemented. IT artifacts were designed and developed by a team of 21 professionals and students with varied technical and artistic expertise. Because this research is supported by a large and diverse group of developers, an ambitious program of design and development was organized, including five components that address specific aspects of the problem space, enabling exploration of our research questions.

Artifact 1: Citizen Sort

Four of the major artifacts of this design effort are organized around a fifth, a portal website (tentatively dubbed *Citizen Sort*) designed to direct participants to a variety of tools and games for biological classification. The portal website controls global functionality, including features like user account management, administrative management of tools and games, content management of the website itself, dissemination of project data, and management of subsidiary projects. A centralized database ties all IT artifacts in this project tightly together.

The IT artifacts hosted on the *Citizen Sort* website include tools and games, organized along a continuum from “tool-like” to “game-like.” Arranging the systems in this manner allows for comparative evaluations of participant motivation with regard to tools, games, and IT artifacts that fall somewhere in between. In addition, this arrangement allows research-

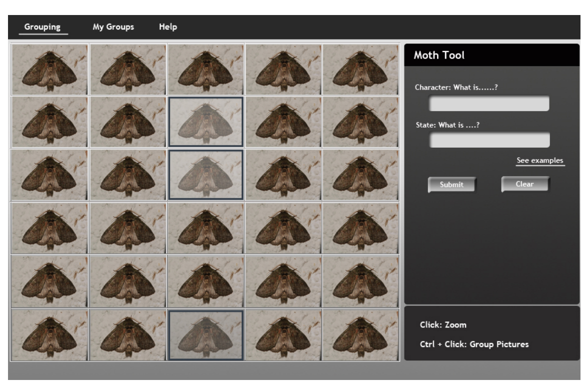

Figure 3: Hunt & Gather *Prototype Interface*

ers to manipulate specific website elements to either direct participants to tools or games or allow participants to self-sort based on their individual interests.

Artifact 2: Hunt & Gather Tool

Hunt & Gather is a "true" tool, designed without additional motivational elements (see [30] for a discussion of motivators vs. satisfiers in web applications). *Hunt & Gather* lets users create characters and states for themselves, tag large numbers of photos with those characters and states, and let other knowledgeable individuals work with the characters, states, and photos on a per project basis. *Hunt & Gather* will allow information science scholars to explore the motivations of users who are attracted to citizen science tools, rather than games; it is hypothesized that these users will be experts or enthusiasts. Furthermore, characters and states created by novices or enthusiasts can be compared to characters and states generated by professional scientists. *Hunt & Gather* will help explore how good non-expert users are at producing characters and states that might be useful to experts in the biological sciences.

Artifact 3: Happy Moths

Happy Moths (to be renamed for each new instantiation: *Happy Sharks, Happy Plants, etc.*) is a "game-like tool," in that it offers tool-like functionality but with some elements of a game. Participants are presented with a set of ten photographs of some organism (in *Happy Moths*, pictures of moths) and then asked to identify the various character-states of each.

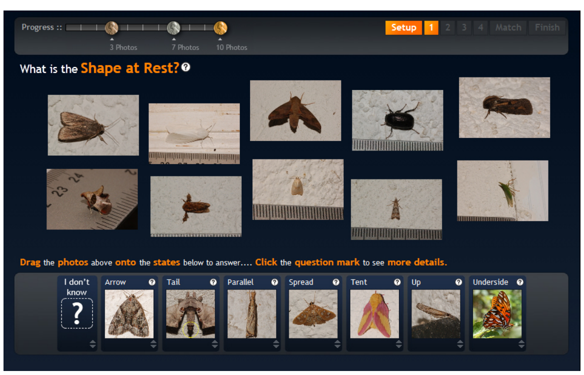

Figure 4: Happy Moths *Prototype Interface*

One difference between *Happy Moths* and *Hunt & Gather* is that the design aims to increase participant motivation by providing a score (per round and overall) giving feedback on performance. *Happy Moths* players are scored based on how well their classification decisions match for a previously classified-photo that is seeded into the game (the "Happy Moth"). Because players will not know which photo is the Happy Moth until the end of each game, they need to do well on all photos to ensure a high score.

A second difference is that *Happy Moths* is built around characters and states established by professional scientists as a useful taxonomic key. *Happy Moths* is a more controlled experience for users, and may ultimately produce more reliable data when used by novices or enthusiasts with limited classification experience. As well, the quality of a player's performance on the Happy Moth can be taken as evidence of their data quality, and agreement among classifications performed by different users on the same photo can be used as an indicator of data validity.

Artifact 4: Happy Moths (Mobile)

Happy Moths (Mobile) will be a mobile version of the *Happy Moths* game, developed as an HTML5 app and deployable on a variety of mobile devices. The mobile version of the game will be very similar to the web-based version of *Happy Moths* (both systems draw upon the same API and database). *Happy Moths (Mobile)* will introduce mobile technology as a variable in comparative evaluation studies; it will be useful in exploring whether mobile technologies make this game seem more or less game-like to users. It can also be used to collect data about where, how, and by whom the mobile version of the game might be used, and it will be possible to compare the quality of data produced by both version of the game.

Artifact 5: Forgotten Island

Finally, an important goal of this research is to explore the full range of the "tool-like" to "game-like" continuum. Few citizen science projects attempt to leverage the power of storytelling or fantasy in games to motivate users. In [14-16], these elements and others are noted as key motivators in educational games; it is hypothesized that such motivators will hold true in citizen science games as well. To explore this hypothesis, as well as to generate insight into the kinds of users who might be attracted by such a game, the fifth IT artifact in this design science project is an adventure game called *Forgotten Island*.

Figure 5: Forgotten Island *Concept Artwork*

Forgotten Island is story driven, featuring an island to explore and a mystery to unravel. Players still classify insects, plants, or animals as in *Happy Moths*, but the classification task is motivated by the story and designed to fit into the background texture of the game. Players use classification as a way to earn game money that can be used to purchase equipment or items to progress the fantasy story.

Forgotten Island allows researchers to explore how endogenous reward systems can motivate players to participate in a scientific collaboration. It will also help researchers explore how established taxonomies of motivational game features for learning [e.g. 14, 15, 16] might apply to non-educational games. Two additional and conflicting hypotheses will be evaluated: 1) That a fantasy adventure game will improve scientific data quality because players will be immersed in the game experience, motivated, and willing to provide high quality data, or 2) That a fantasy adventure game will reduce data quality because players will be more interested in progressing the story than in doing science, and will be willing to "cheat" on the science task to get ahead in the game.

Evaluation Method

Prior to the start of system development, background research was conducted in the form of literature review, analysis of ongoing citizen science project systems, and SME interviews. Ten SME interviews with nine scientists and developers who are currently undertaking citizen-science projects were conducted. This phase of the project informed research questions and planning for the IT artifacts to be developed. As design progressed, additional SMEs were consulted, including naturalists with expertise in classification. Consultation with these and other experts is ongoing, shifting between formal, interview-style consultation and informal participatory-research approaches [6]. This research is in the design stage, with limited formal evaluation so far. One formal focus group evaluation session brought four expert entomologists together codify their knowledge of the classification task and to collect their impressions of the *Happy Moths* game. The design process itself has also served as an important source of data.

Preliminary Results and Discussion

Participant Groups

During the *Happy Moths* focus group session, SMEs helped to define three groups of potential participants who will be important for this research: 1) experts (professional scientists), 2) enthusiasts (individuals with intrinsic interest in science and/or the particular topic of a citizen science project), and 3) gamers (ordinary citizens with no particular interest in citizen science, but an interest in online games or entertainment). Because it may be difficult for some projects to attract enough expert and enthusiast users to be viable, the gamer user group is of particular interest. The gamer group is hypothesized to be much larger than the enthusiast or expert groups, making it a potentially valuable source of participants. However, the gamer group, by definition, is composed of individuals who have virtually no knowledge of scientific classification; finding ways to make the classification task enjoyable and, critically, understandable to these users will be an important outcome. One way of addressing this challenge, used in *Happy Moths*, is to have SMEs generate character questions and state answers that make sense to laypeople. So, for example, *Happy Moths* asks about simpler character-state combinations such as color or shape, and avoids complex questions about "discal spots," orbicular spots," "reniform spots," etc. In many cases, technical language has also been simplified to help lay users understand characters and states without the need for extensive training. In the *Happy Moths* focus group, SMEs had conflicting opinions about these approaches; some agreed that simplifying the tasks and language would be beneficial and still produce good data, while others felt that more technical nomenclature should be preserved as a learning opportunity for gamers.

This disagreement raises another point about the differences between users: systems that motivate gamers may actually be de-motivating to enthusiasts and vice-versa. In the focus group session, researchers suggested that systems designed to appeal to gamers (e.g., *Forgotten Island*) have a high likelihood of alienating enthusiasts. Enthusiasts are seeking opportunities to explore their passions and interests, while gamers are seeking entertainment. Over the course of design and evaluation so far, it has emerged that as a collective intelligence game focuses more on entertainment, it imposes increasing obstacles on enthusiasts who seek rapid access to their hobby of choice. For example, *Forgotten Island* paces the classification task and requires players to explore a variety of locations, collect items, and undertake many other story-driven activities besides classification. For an enthusiast, interested in classification, these extra activities may be perceived as annoying wastes of time, rather than as fun. Similarly, SMEs frequently suggest that players will be more engaged and motivated if they learn something about science, but it is not clear that gamers will be similarly motivated.

The Role of Iteration

The purpose of taking each project in this design science study through several design iterations is threefold: each iteration 1) improves the IT artifact's ability to address the problem space, 2) produces new research findings, and 3) helps to eliminate poor system design as a confounding factor for research.

In the case of *Citizen Sort*, many specific design decisions have been discussed with the project's SMEs, particularly the decision-making that went into the *Happy Moths* game, which has (because it best encapsulates the classification task) received the most formal evaluation to date. Many design decisions have been upheld, while a few have been questioned (e.g., the visual style of *Happy Moths*, where expert reviewers suggested that a more "natural" or "nature-themed" design would better appeal to enthusiast users). In some cases, design decisions have been rejected outright. In the first iteration of *Happy Moths*, music was included, but focus group SMEs and the developers themselves unanimously rejected the choice to include music after testing it in several different settings. Now entering its third iteration, *Happy Moths* has no music and a streamlined game mechanic that is expected to be more fun for players.

Task Gamification vs. Task Incorporation

Collective intelligence and citizen science games are often developed by "gamifying" a task. The *Happy Moths* game adopts this approach, taking a classification task and adding game elements to it: a game-like visual design, scores for doing well, achievements for long-term involvement, and leader boards and high scores to promote competition between players.

An alternative approach is to make the task just one small piece of a larger game experience. To be effective at generating data, the task must be incorporated in a way that makes it critical to progress through the game. This approach is rarely pursued, possibly because of its scope and difficulty; developing a fantasy/story game like *Forgotten Island* is an exponentially larger effort than "gamifying" a scientific task. By placing the scientific task into the background of a fantasy game, developers are suddenly confronted with a host of new activities: writing the story, designing locations, creating characters, sound design, puzzle generation, and much more. Making the scientific task seem sensible in the context of a fantasy story is a particularly difficult writing challenge.

A third approach, not part of this research design, but offering interesting possibilities for future study, is to turn a scientific or collective intelligence task into a form of "payment" for play. Many casual games have successfully adopted a model where micro-payments unlock game items, new content, or new levels. Substituting classification for cash payment could be an effective way to reward users for their help and attract gamers to a project.

Friction

One complexity of the design science approach is the friction that generates through competition between problem space, research goals, and feasibility to develop the IT artifact. These factors each require tradeoffs among the others. In the *Citizen Sort* project, SMEs want to take ownership of a suite of games and tools to support a citizen classification effort. Their primary goal is that these should produce large amounts of very high quality data. Virtually all other considerations are secondary. From an information science research perspective, however, the interest is in how different kinds of games or tools can motivate different kinds of users and produce different qualities of data. It matters less that each individual tool or game produce the best quality data or attract the right kind of users, than that each game or tool helps generate useful knowledge about the research questions of interest. This means that a game like *Forgotten Island* could produce extremely poor classification data but still be a research success in providing evidence of cheating effects or other flaws in the fantasy/story approach. This outcome would, of course, be considered a failure by SMEs.

In [20], the need for multi-disciplinary expertise as well as expert developers on a design science project is noted, the better to adequately address both the problem space and research goals. Galison [9] describes how such collaborations can be difficult when friction between the varying goals of different interested parties develops. Galison describes the idea of "trading zones" [9] to accommodate the needs of various collaborators through a negotiating process. *Citizen Sort's* project manager takes a central role in these negotiations, coordinating various groups of SMEs and developers, ensuring that natural science and information science requirements are balanced, and verifying that the project scope is feasible for the development team. Evaluation efforts have validated "trading zone" efforts on this project, with research goals and the problem space largely complementing rather than conflicting with each other.

CONCLUSION

Design science is an approach to scientific inquiry where research goals are pursued through the development of an IT artifact positioned to address a real-world problem. This approach has many strengths, including the ability to tightly control research efforts while still enacting them within realistic use contexts. In addition, evaluation of design science efforts can address numerous research questions.

One constraint of design science is the friction that can develop between research goals, the problem space, and system feasibility. While good project management and careful attention to both researcher and stakeholder needs can mitigate these effects, friction is virtually impossible to eliminate entirely.

Nonetheless, as the *Citizen Sort* project demonstrates, design science can be a valuable approach to exploring design issues in citizen science, purposeful gaming, and collective intelligence.

REFERENCES

1. Bradford, B.M. and Israel, G.D., *Evaluating Volunteer Motivation for Sea Turtle Conservation in Florida*. 2004, University of Florida, Agriculture Education and Communication Department, Institute of Agriculture and Food Sciences: Gainesville, FL. p. 372.
2. Brooks, F.P., Jr., No Silver Bullet: Essence and Accidents of Software Engineering. *IEEE Computer*, 1987. **20**(4): p. 10-19.
3. Brooks, F.P., Jr., The Computer Scientist as Toolsmith II. *Communications of the ACM*, 1996. **39**(3): p. 61-68.
4. Cohn, J.P., Citizen Science: Can Volunteers Do Real Research? *BioScience*, 2008. **58**(3): p. 192-107.
5. Crowston, K. and Fagnot, I. The motivational arc of massive virtual collaboration. in *IFIP WG 9.5, Working Conference on Virtuality and Society: Massive Virtual Communities*. 2008. Lüneberg, Germany.
6. DeWalt, K.M. and DeWalt, B.R., *Participant Observation: A Guide for Fieldworkers*. 2002, Walnut Creek, CA: AltaMira Press. 287.
7. Dix, A., Human–computer interaction: A stable discipline, a nascent science, and the growth of the long tail. *Interacting with Computers*, 2010. **22**: p. 13-27.
8. Ellis, G. and Dix, A., *An explorative analysis of user evaluation studies in information visualisation*, in *Proceedings of the 2006 AVI workshop on BEyond time and errors: novel evaluation methods for information visualization*. 2006, ACM: Venice, Italy.
9. Galison, P., *Image and Logic: A Material Culture of Microphysics*. 1997, Chicago: The University of Chicago Press.
10. Gregor, S., The Nature of Theory in Information Systems. *MIS Quarterly*, 2006. **30**(3).
11. Hevner, A.R., March, S.T., Park, J., and Ram, S., Design Science in Information Systems Research. *MIS Quarterly*, 2004. **28**(1): p. 75-105.
12. King, K. and Lynch, C.V., The Motivation of Volunteers in the nature Conservancy - Ohio Chapter, a Non-Profit Environmental Organization. *Journal of Volunteer Administration*, 1998. **16**(5).
13. Law, E. and von Ahn, L., *Input-agreement: a new mechanism for collecting data using human computation games*, in *Proceedings of the 27th international conference on Human factors in computing systems*. 2009, ACM: Boston, MA, USA.
14. Malone, T., W., *What makes things fun to learn? heuristics for designing instructional computer games*, in *Proceedings of the 3rd ACM SIGSMALL symposium and the first SIGPC symposium on Small systems*. 1980, ACM: Palo Alto, California, United States.
15. Malone, T.W., *Heuristics for designing enjoyable user interfaces: Lessons from computer games*, in *Proceedings of the 1982 conference on Human factors in computing systems*. 1982, ACM: Gaithersburg, Maryland, United States.
16. Malone, T.W. and Lepper, M., *Making learning fun: A taxonomy of intrinsic motivations for learning*. Aptitude, learning, and instruction: Vol. 3. Cognitive and affective process analyses, ed. R. Snow and M. Fair. 1987, Hills-Dale, NJ: Erlbaum.
17. Malone, T.W., Laubacher, R. and Dellarocas, C.N., *Harnessing Crowds: Mapping the Genome of Collective Intelligence*, in *MIT Sloan Research Paper No. 4732-09*. 2009.
18. March, S.T. and Smith, G.F., Design and natural science research on information technology. *Decision Support Systems*, 1995. **15**(4): p. 251-266.
19. Peffers, K., Tuunanen, T., Rothenberger, M., and Chatterjee, S., A Design Science Research Methodology for Information Systems Research. *Journal of Management Information Systems*, 2007. **24**(3): p. 45-77.
20. Prestopnik, N., Theory, Design and Evaluation – (Don't Just) Pick any Two. *AIS Transactions on Human-Computer Interaction*, 2010. **2**(3): p. 167-177.
21. Raddick, M.J., et al., *Citizen science: status and research directions for the coming decade*, in *AGB Stars and Related Phenomenastro 2010: The Astronomy and Astrophysics Decadal Survey*. 2009. p. 46P.
22. Raddick, M.J., et al., Galaxy Zoo: Exploring the Motivations of Citizen Science Volunteers. *Astronomy Education Review*, 2010. **9**(1): p. 010103-18.
23. Rittel, H.J. and Webber, M.M., *Planning Problems Are Wicked Problems*, in *Developments in Design Methodology*, N. Cross, Editor. 1984, John Wiley & Sons: New York.
24. Smith, K.M., Producing the Rigorous Design Case. *International Journal of Designs for Learning*, 2010. **1**(1).
25. von Ahn, L., et al., *Improving accessibility of the web with a computer game*, in *Proceedings of the SIGCHI conference on Human Factors in computing systems*. 2006, ACM: Montreal, Quebec, Canada.
26. von Ahn, L., *Human computation*, in *Proceedings of the 4th international conference on Knowledge capture*. 2007, ACM: Whistler, BC, Canada.
27. von Ahn, L., *Human computation*, in *Proceedings of the 46th Annual Design Automation Conference*. 2009, ACM: San Francisco, California.
28. Wiggins, A. and Crowston, K., Developing a conceptual model of virtual organizations for citizen science. *International Journal of Organizational Design and Engineering*, 2010. **1**(1/2): p. 148-162.
29. Wiggins, A. and Crowston, K. From Conservation to Crowdsourcing: A Typology of Citizen Science. in *44th Hawaii International Conference on System Sciences*. 2011. Kauai, Hawaii.
30. Zhang, P. and von Dran, G.M., Satisfiers and dissatisfiers: A two-factor model for Website design and evaluation. *Journal of the American Society for Information Science and Technology*, 2000. **51**(14): p. 1253.